

DTIC File Copy

AFHRL-TP-89-34

2

AIR FORCE



AD-A212 772

**HUMAN
RESOURCES**

**IDENTIFICATION AND EVALUATION OF METHODS
TO DETERMINE ABILITY REQUIREMENTS FOR
AIR FORCE OCCUPATIONAL SPECIALTIES**

Walter E. Driskill

The Maxima Corporation
8301 Broadway, Suite 212
San Antonio, Texas 78209

Johnny J. Weissmuller
Dwight C. Hageman

Metrica, Incorporated
8301 Broadway, Suite 215
San Antonio, Texas 78209

Linda E. Barrett

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

August 1989

Interim Technical Paper for Period October 1988 - June 1989

Approved for public release; distribution is unlimited.

DTIC
ELECTE
SEP 25 1989
E

LABORATORY

**AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601**

89 9 25 083

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Division

DANIEL L. LEIGHTON, Colonel, USAF
Chief, Manpower and Personnel Division

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) AFHRL-TP-89-34			
6a. NAME OF PERFORMING ORGANIZATION Metrica, Incorporated		6b. OFFICE SYMBOL (if applicable)		7a. NAME OF MONITORING ORGANIZATION Manpower and Personnel Division		
6c. ADDRESS (City, State, and ZIP Code) 8301 Broadway, Suite 215 San Antonio, Texas 78209			7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235-5601			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory		8b. OFFICE SYMBOL (if applicable) HQ AFHRL		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F41689-88-D-0251-5014		
3c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601			10. SOURCE OF FUNDING NUMBERS			
			PROGRAM ELEMENT NO. 62703F	PROJECT NO. 7719	TASK NO. 18	WORK UNIT ACCESSION NO. 61
11. TITLE (Include Security Classification) Identification and Evaluation of Methods to Determine Ability Requirements for Air Force Occupational Specialties						
12. PERSONAL AUTHOR(S) Driskill, W.E.; Weissmuller, J.J.; Hageman, D.C.; Barrett, L.E.						
13a. TYPE OF REPORT Interim		13b. TIME COVERED FROM Oct 88 TO Jun 89		14. DATE OF REPORT (Year, Month, Day) August 1989		15. PAGE COUNT 72
16. SUPPLEMENTARY NOTATION <i>to look</i>						
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)			
FIELD	GROUP	SUB-GROUP	ability requirements → personnel classification, taxonomy			
05	08		→ Air Force personnel, personnel selection			
05	09		CODAP task analysis			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) This effort reviewed the various methods which have been developed for the determination of ability requirements from job analysis data. A review of the literature identified 36 general and special purpose taxonomies which have been applied to the description of job and/or worker characteristics. Seven methods were chosen for further evaluation: Functional Job Analysis, Job Element Method, Position Analysis Questionnaire, Occupation Analysis Inventory, General Work Inventory, Threshold Traits Analysis System, and Ability Requirements Scales. These approaches were further evaluated on the basis of (a) effectiveness for the purpose of selection, (b) compatibility with Air Force occupational task analysis data, and (c) utility for Air Force operational use. It was determined that the Ability Requirements Scales (ARS) method developed by Fleishman was most appropriate for the identification of ability requirements of different Air Force occupational specialties. The ARS approach was most compatible with Air Force occupational task analysis data, provided the most extensive coverage of abilities, and provided the most direct link with tests of specific abilities. However, the ARS taxonomy appears to be incomplete with regard to all abilities relevant to job performance. Development of an Air Force taxonomy of occupational abilities was recommended. Ability						
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified			
22a. NAME OF RESPONSIBLE INDIVIDUAL Esther M. Barlow			22b. TELEPHONE (Include Area Code) (512) 536-2876		22c. OFFICE SYMBOL AFHRL/SCVE	

Item 19 (Concluded):

requirements for an occupational specialty would be determined through subject-matter expert ratings of ability requirements for task statements categorized by the verbs in the task statements. ~~This approach~~ would enable generalization across occupational specialties. If this method were applied and ratings established for all relevant verbs, ability requirement profiles could be generated for any Air Force occupational specialty without the expense of subject-matter expert conferences. *See serial 12*

FEB '8

IDENTIFICATION AND EVALUATION OF METHODS
TO DETERMINE ABILITY REQUIREMENTS FOR
AIR FORCE OCCUPATIONAL SPECIALTIES

Walter E. Driskill

The Maxima Corporation
8301 Broadway, Suite 212
San Antonio, Texas 78209

Johnny J. Weissmuller
Dwight C. Hageman

Metrica, Incorporated
8301 Broadway, Suite 215
San Antonio, Texas 78209

Linda E. Barrett

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

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

Reviewed by

Jacobina Skinner
Acting Chief, Officer Selection and Classification Function

Submitted for publication by

Lonnie D. Valentine, Jr.
Chief, Force Acquisition Branch

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

The United States Air Force maintains detailed task analysis data which provides extensive information pertaining to the actual tasks performed within enlisted and officer occupational specialties. From this data base the tasks which are most typical and important can be identified through examination of the proportion of job incumbents who perform a task, the proportion of time spent on the task, and ratings of task difficulty and/or importance to a training curriculum. Task analysis data have been utilized primarily for classification of jobs and determination of training requirements. Task analysis data can also contribute to the selection process once the abilities required to perform similar tasks are identified. Determination of ability requirements from task analysis data would enable the generation of ability profiles for different occupational specialties and would provide empirical justification for selection tests of abilities required across or within occupational specialties.

This effort reviewed the various methods which have been developed for the determination of ability requirements from job analysis data. A review of the literature identified 36 general and special purpose taxonomies which have been applied to the description of job and/or worker characteristics. Seven methods were chosen for further evaluation: Functional Job Analysis, Job Element Method, Position Analysis Questionnaire, Occupational Analysis Inventory, General Work Inventory, Threshold Traits Analysis System, and Ability Requirements Scales. The first five methods provide descriptive information about work requirements from which ability requirements may be determined. The last two are ability requirements methods which identify ability demands from task or job descriptive data.

These approaches were further evaluated on the basis of (a) use of a specified taxonomy of abilities, (b) comprehensiveness of the ability taxonomy, (c) reported reliability and validity, (d) level of job analysis (task or whole job), (e) procedure used to determine ability requirements, (f) linkage of abilities to existing tests, (g) compatibility with Air Force task analysis data, and (h) ease of use. It was determined that the Ability Requirements Scales (ARS) developed by Fleishman were most appropriate for the identification of ability requirements of different Air Force occupational specialties. The ARS approach was most compatible with Air Force occupational

task analysis data, provided the most extensive coverage of abilities, and provided the most direct link with tests of specific abilities.

While the ARS were evaluated as most effective of the established methods, there is reason to believe that the scales are insufficient to cover the requirements of all Air Force occupational tasks. Specifically, communication and interpersonal skills are not adequately covered. Also, cognitive ability research using computer testing technology should be investigated for additional relevant cognitive ability constructs. It was recommended that the ARS be refined to include any additional constructs of interest, and delete or consolidate existing constructs in order to tailor the taxonomy to Air Force occupational requirements.

Ability requirement, for an occupational specialty would be determined through subject-matter expert ratings of ability requirements for task statements categorized by the verbs in the task statements. This approach would enable generalization across occupational specialties. If this method were applied and ratings established for all relevant verbs, ability requirement profiles could be generated for any Air Force occupational specialty without the expense of subject-matter expert conferences.

PREFACE

This work was completed under Task 771918, Selection and Classification Technologies, which is part of a larger effort in Force Acquisition and Distribution. It was subsumed under work unit number 77191861, The Identification, Development, and Validation of New Predictors of Air Force Officer Success. This effort was initiated to investigate the feasibility of utilizing job analysis data for the identification of abilities and/or aptitudes which (a) underlie successful performance across and within Air Force occupational specialties, and (b) are relevant for the purpose of selection. The authors wish to thank Dr. Bruce Gould, Mr. Wayne Archer, and Mr. William Phalen for their interest, encouragement, and insightful comments.

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
I. INTRODUCTION.....	1
Purpose.....	1
Background.....	2
Approach.....	3
II. LITERATURE REVIEW.....	4
Introduction.....	4
Issues Related to Selection.....	4
Criteria for Evaluating Taxonomies.....	5
Conceptual Basis.....	5
Special Purpose Taxonomies.....	20
Rating Scales: Measurement and Procedural Issues.....	23
Internal Validity: Interrater Reliability and Taxonomic Adequacy.....	26
External Validity: Job Component and Criterion-related Validity.....	28
Utility.....	31
Summary and Conclusions from the Literature Review.....	32
III. EVALUATION OF TAXONOMIES FOR AIR FORCE USE.....	34
CODAP Compatibility.....	35
Linkage with Construct-Valid Selection Instruments.....	35
Ease of Use.....	37
Adequacy of Coverage of Ability Requirements.....	38
Generalizing Ability Requirements Across Specialties.....	40
Descriptive Data for Specialties for Which OSM Do Not Exist.....	42
Summary of Evaluation.....	43
Evaluation Factors.....	43
Evaluation Summary.....	43
Conclusions.....	44
Recommendations.....	47
REFERENCES.....	49
APPENDIX: DETAILED EVALUATIVE DATA.....	55

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	General Work Inventory (GWI) Content.....	10
2	Ability Requirement Scales.....	17
3	An Example of Ability Requirements Scales Definition.....	18
4	Cognitive Abilities.....	19
5	Classes of Candidate Taxonomies.....	34
6	Evaluation Factors.....	44
7	Summary of Evaluation Data.....	45

I. INTRODUCTION

Purpose

This effort was initiated to review alternative procedures for identifying ability requirements of Air Force occupational tasks. The United States Air Force Occupational Measurement Center (USAF/OMC) collects and maintains task analysis data which describe the specific tasks performed by enlisted and officer personnel in most Air Force occupational fields. Task-related data include ratings of whether the task was performed by an incumbent and what proportion of time is spent on the task. Ratings of task difficulty and/or training priority are collected for certain occupational specialties. These data have been utilized primarily for job classification and determination of training curricula (Mitchell, Ruick, & Driskill, 1988). The potential contribution of task-related data for the enhancement of selection decisions has not been fully explored.

There are two basic approaches by which task information can be utilized in the selection decision process. The most straightforward method is that of job content validity, as described in the Uniform Guidelines (1978). Content validity begins with the identification of all important representative tasks of the job. These tasks then make up the content of the selection test(s). Thus an applicant for an automotive mechanic position may be requested to diagnose and repair a typical automotive problem.

The job content approach to selection test development is not applicable to Air Force occupations for two main reasons. The immense variety of Air Force occupational tasks and the large number of applicants prohibit the development of job content-oriented selection measures. Another constraint to this approach is the extensive training that most Air Force selectees experience. Procedural knowledge of the majority of tasks is gained through training, and thus is not required at the time of selection.

The second approach to the utilization of task information for selection also begins with the identification of important representative tasks. The tasks are then examined to identify the abilities that are required to learn or perform the task successfully. Once the prerequisite abilities are identified, tests are chosen or developed to measure the desirable abilities. Tests may consist of biographical data, responses to situational items, and/or written tests of specific abilities.

Identification of ability requirements of similar tasks would enable the identification of the ability requirements of a specific occupation or any group of occupations. Ability profiles could be used to categorize occupational specialties according to ability requirements. Once ratings of ability requirements were collected for a common pool of tasks, ability profiles could be generated for a restructured occupational specialty.

Background

The Air Force Comprehensive Occupational Data Analysis Programs (CODAP) enable the analysis of detailed task inventory data gathered from job incumbents of each Air Force occupational specialty. These data provide detailed information as to what military personnel in a particular specialty actually do. Task inventory profiles describe the percentage of incumbents who perform particular tasks and the relative amount of time spent on each task. The data base is used for many purposes including job classification, assignment, and training decisions.

Training requirements are identified through the collection of additional task information, such as ratings of task difficulty and the extent to which the task should be emphasized in training (Mitchell, Ruck, & Driskill, 1988). While this additional task information is useful in the identification of aptitudes and abilities to be assessed for selection, it is not sufficient. Applicants can be selected on the basis of aptitudes required for success in training, but not on the basis of knowledge or abilities which can be expected to be acquired during formal or on-the-job training (Uniform Guidelines, 1978). Further, occupational duties may require additional aptitudes or abilities on the job that were not identified as training requirements, or that are not included in the training curriculum, such as interpersonal skills. Thus, while the training requirements data are relevant to this effort, they are not sufficient for the identification of aptitudes which should be considered for selection decisions.

There have been previous efforts to identify the abilities required for different Air Force occupational specialties. Most efforts were not based on task inventory-CODAP analysis or were not oriented for the purpose of selection. These efforts have either administered a separate survey, such as the Position Analysis Questionnaire (PAQ), to an occupational group of interest, or they generated a list of abilities based on subject-matter expert (SME) perceptions of the whole job. This 'whole job' approach is not sufficient to establish either content or construct validity. In order to meet requirements for content or construct validity the ability requirements must be derived from task analysis data (Uniform Guidelines, 1978). Also, previous efforts usually did not take full advantage of CODAP resources. An Air Force study which did utilize a predefined taxonomy of abilities and determined requirements from task data succeeded in identifying the perceptual and psychomotor requirements of 35 Air Force specialties (Siegel, Federman, & Welsand, 1980).

There are numerous methods which have been developed to ascertain aptitude and ability requirements from task inventory data. These methods differ in procedure and purpose. For example, several methods have a predetermined taxonomy of job-related abilities (Fleishman, 1988; Lopez, 1988; McCormick & Jeanneret, 1988; Siegel, Federman, & Welsand, 1980); other methods generate a taxonomy of abilities as a result of task inventory data analysis (Fine, 1988; Primoff & Eyde, 1988). Methods also differ in how the required type and level of abilities are identified. Abilities may be derived automatically from an

existing system which links task verbs to ability requirements (Fine, 1988) or may be established through interrater agreement using SMEs, job incumbents, and/or supervisors.

In addition, ratings may be directed toward a variety of task characteristics, such as importance, frequency, practicality for selection, degree of discrimination between poor and superior workers, or consequences of incompetence. Methods also differ in how ratings are used for the purpose of selection. Some methods have a particular algorithm which results in an index which describes the importance of the task or ability for selection (Primoff & Eyde, 1988); other methods may simply rank the order of the ratings.

Approach

In this study, the diversity of established procedures were fully explored via a review of the literature (Section II) to identify the most efficient system for Air Force selection purposes. The differences in procedures were evaluated for (a) effectiveness in general (reliability and degree of content validity), (b) effectiveness for the purpose of selection (consideration of issues pertaining to selection decisions), (c) compatibility with CODAP data bases, and (d) projected cost of implementation. Evaluation factors were identified, and an evaluation plan was developed for the determination of the utility of the procedures when applied to Air Force selection (Section III).

II. LITERATURE REVIEW

Introduction

The literature on taxonomies of human performance is extensive. Taxonomies, for example, have been developed for such diverse uses as training and learning (Gagne, 1977; Lumsdaine, 1960); acceleration stress behaviors (Chambers, 1963), human-machine interfacing (Finley, Obermayer, Bertone, & Muckler, 1970); selection and classification (Fleishman & Quaintance, 1984); theory development, such as in the taxonomies developed for information theory research (Levine & Teichner, 1973; Posner, 1964); and leadership (Fleishman, 1972) as well as for many other purposes. Generally, these taxonomies have ranged from those designed for a specific purpose to general-purpose taxonomies. The more specific taxonomies have limited usefulness for other applications. More general purpose taxonomies, such as Fleishman's taxonomy of ability requirements (Fleishman & Quaintance, 1984), have greater utility, since they may be applied across multiple domains. As Fleishman and Quaintance observe, despite the multitude of taxonomies available, few have broad application across the study of human performance.

This literature review is directed at defining the criteria by which the multitude of taxonomies could be evaluated and at preparing a list of candidate taxonomies. It is based on the consideration of 36 taxonomies from which seven are selected as having broad application and potential for use in identifying the ability requirements of Air Force officer and airman specialties. These seven methodologies are examined in light of the criteria for the development, validation, and use of a taxonomy.

Issues Related to Selection

The first step in identifying a candidate taxonomy is a clear and unambiguous definition of the purpose the taxonomy is to serve and the purpose in this case is selection. While this purpose seems clear and unambiguous, identification of candidate taxonomic approaches is nevertheless complicated by the train of events following selection that affect the kind of taxonomy employed (Peterson & Bownas, 1982). If, for example, no postemployment training is to be provided, a taxonomy that includes knowledge of job-related subject-matter procedures is most useful. An appropriate selection approach would be the use of work sample, job knowledge, or similar tests to evaluate applicants' potential for doing the tasks in each category without training; testing for job content mastery is desirable. In such a case, a content-validity selection strategy is clearly appropriate.

On the other hand, where there is postselection training, the consensus reflected in the literature (e. g., Fleishman & Mumford, 1988; Fleishman & Quaintance, 1984; McCormick, 1979; Petersen & Bownas, 1982) is that it is more

appropriate to screen and select applicants on the basis of job-related abilities or aptitudes as opposed to screening for job content mastery. These abilities or attributes are generally conceived of as the relatively enduring attributes of the individual that influence a broad range of task performance (Fleishman & Mumford, 1988; McCormick, Mecham, & Jeanneret, 1977), as opposed to the identification of knowledges or skills that may be acquired through training or experience. When aptitudes or abilities are of interest, as opposed to job knowledge and skills, then construct referenced strategies are more appropriate. Thus, in this effort, the main focus is the review of methods utilized to identify the prerequisite ability constructs from task analysis data.

For this effort, general purpose taxonomies and job analysis systems of two kinds are reviewed. First, methodologies for describing jobs in terms of work activities (which include descriptors of a number of mental activities) and technical job content at a generic level from which "abilities" are then derived and are reported. Second, methodologies which derive "abilities" directly from analysis of the specific tasks incumbents perform in their jobs are considered. The latter taxonomies provide no technical content of jobs. These methods are reviewed in the following section.

Criteria for Evaluating Taxonomies

Several criteria for evaluating taxonomies for human task performance are apparent in the literature (Fleishman & Quaintance, 1984; Peterson & Bownas, 1982). The criteria suggested in the literature include (a) conceptual basis of the taxonomy; (b) qualitative v. quantitative classification; (c) internal and external validity; and (d) utility. In the following sections, general purpose job analysis and taxonomic methodologies are critically reviewed on the four criteria. A description of the job analysis program of the USAF/OMC is also included.

Conceptual Basis

Conceptual basis refers to the type of components described within a taxonomy and is the major factor differentiating the general purpose taxonomies. In regard to the various conceptual approaches, Fleishman and Mumford (1988) characterize four approaches to the classification of human task performance: behavioral description, behavioral requirements, ability requirements, and task characteristics. Each of these approaches, at least in some of their variations, can lead to the identification of human attributes that facilitate task performance.

In the selection of a taxonomy for Air Force use, the conceptual approach employed by a taxonomy bears directly on three crucial issues to be considered in Section III. These issues pertain to the compatibility of the taxonomic

methodology with existing Air Force job analysis procedures, the linkage of job descriptive requirements to ability requirements, and the application of job component validity. Elaboration of the importance of these factors for the Air Force are provided in Section III. The four basic approaches and various methods are described and discussed in the following section.

1. The Behavioral Description Approach, as its name implies, consists of a specification of what workers actually do while performing their jobs. This approach produces a general-purpose analysis of jobs. The products of the application of descriptive methods range from highly task-specific to much more general task information. Usually emphasis is placed on the overt behaviors of workers or certain subjective terms which describe what the worker does. Several systems of job analysis representing this approach are in use, each differing from the others with respect to the level and the context of their descriptors. The following behavioral description approaches will be discussed in this section: (a) USAF Occupational Survey Methodology, (b) Position Analysis Questionnaire, (c) Occupational Analysis Inventory, (d) General Work Inventory, (e) Job Element Method, and f) Functional Job Analysis.

The Occupational Survey Methodology (OSM) is a job analysis methodology, which provides the most specific description of the work that job incumbents perform. The Occupational Measurement Data Base provides a rich source of specific data about tasks, jobs, and Air Force specialty (AFS) structure. It does not provide information about the ability requirements of tasks or jobs---these must be derived from some other taxonomy.

The OSM, of all of the descriptive approaches, provides the broadest and clearest data about the tasks workers perform. More importantly, this methodology provides information about the structure of jobs that other methods do not permit. Other methods sample from jobs as a whole. In such broad sampling the SMEs are usually requested to indicate the "most important tasks" in the occupation, and the variance of jobs within an occupation can easily be missed. This variance should not be ignored as AFSs are structured to accommodate career progression, meaningfulness of individual jobs, and promotion potential (Driskill & Mitchell, 1979). It is in the variant jobs that the real essence of an AFS may be captured. The OSM, as will be indicated below, provides for the identification of the variance within an AFS.

OSM data are central to many classification and training decisions, serving as a principal basis upon which personnel functions that follow selection operate. Also, several personnel research activities, such as aptitude requirements, Training Decision System, and job performance measurement, are based on use of OSM data (Mitchell & Driskill, 1986). Also, several personnel research activities, such as aptitude requirements, Training Decision System, and job performance measurement, are based on use of OSM data (Mitchell, Ruck, & Driskill, 1988). The OSM is founded on an extensive and continuing research program as well as over 20 years of operational application

in the Department of Defense agencies and in the civilian sector that demonstrate its usefulness and applicability for describing tasks and jobs.

The basic instrument of the OSM, which distinguishes it from most other job analysis systems, is a USAF Job Inventory developed by specialists who employ extensive interviews with SMEs in the AFS being described. The Job Inventory consists of two basic parts: a comprehensive listing of the tasks required to be performed by a specialty and a background information section. The task list is constructed through interviews with incumbents of the specialty. The interview process continues until, in judgment of the SME and a review panel, the AFS is fully described. Task development and description are based on specific guidelines having to do with the following criteria: differentiate among skill level of workers, independence, time ratability, and language consistent with worker usage (Driskill, Weissmuller, & Staley, 1987). A typical task statement would read as: "Remove or replace transmitters," or "Type letters or memoranda." The background information section elicits demographic information as well as work environment data, such as equipment used or maintained, from job incumbents.

Results of administration of the Job Inventory to job incumbents provide, then, basic descriptive data about tasks, demographic information about job incumbents, and descriptive information about the work environment. It does not provide any information about the ability, psychomotor or perceptual skill, knowledge, or physical requirements, although there have been two or three instances in which an OSM survey attempted to elicit information about the physical demands of an AFS.

Upon final validation by SME, the final USAF Job Inventory is printed and distributed to Consolidated Base Personnel Offices worldwide which monitor administration to most of the incumbents (70% to 90%) of an AFS. Inventories are AFS-specific, although in some cases Inventories include two or more AFSs whose incumbents perform related work. The resulting incumbent data are not generalizable to the work of other AFSs, except where two or more are included in a single inventory administration. In responding, each incumbent indicates the tasks he or she performs and rates the relative time spent performing each task relative to each other task.

Job incumbent responses are analyzed, using the Comprehensive Data Analysis Programs (CODAP), for their implication for manpower, personnel, and training programs. CODAP is a system of over 100 programs that provides numerous analytic, summary, and display modes. At the most basic level are job descriptions for individuals or groups of individuals that reflect the tasks performed and the relative percent time incumbents spend performing each of the tasks.

Other analysis programs hierarchically or nonhierarchically cluster tasks. The clustering algorithms are based either on the similarity of the tasks

performed by people and the similarity of the time they spend on each of the tasks, or on coperformance. The first clustering produces clusters of people, or job types, which reflect the structure of the AFS in terms of the similarity of jobs. The second clustering solution results in the formation of clusters of tasks based on the probability of their coperformance. This solution also shows work structure, but from the perspective of which tasks are performed together; it is especially useful for constructing task modules for use in training or other purposes where coperformed tasks are desired.

In addition to the analysis of incumbent task responses, the OSM includes methods for collecting from a small sample of SME judgments about the difficulty of each task and the emphasis that should be given to training the tasks for first-term job performance. These ratings are available for enlisted specialties only. Each SME rates each task on a 9-point difficulty and training emphasis scales. Difficulty is defined as the time it takes job incumbents to learn to do the task proficiently. Interrater agreement is assessed through a program entitled GRPREL which computes and reports the intraclass correlation. These ratings are a basic variable in the computation of job difficulty and for the determination of aptitude requirements of AFSs based on time to learn to perform tasks (Dittmar, Driskill, & Weissmuller, 1987).

Reliability of the percent performing and percent time spent data is high. Christal (1971) randomly split the samples for 10 specialties and computed the correlations of percent performing and percent time spent vectors for each of the specialties. The correlations ranged from .931 to .997. Interrater agreement for task difficulty and training emphasis ratings exceeds .90.

Occupational surveys are applied to both enlisted and officer AFSs. To date, most of the enlisted specialties have been described with the exception of a few very small population specialties, such as musicians. Only about 20 of 45 officer utilization fields have been surveyed (USAF Program Technical Training, 1988). Those not surveyed include most of the pilot fields, medical and legal officers, and engineers, although a recent survey collected responses about pilot tasks from pilots of 56 aircraft. The list of tasks in the Job Inventory appears to be exhaustive, representing most if not all of the Air Force pilot technical tasks. Management-related tasks were not included in the pilot survey.

In addition, there have been two surveys of leadership, management and supervision, and communication tasks across and within Air Force occupational specialties. No technical task performance is described by these surveys. In the officer fields, the total sample of grades 0-1 through 0-6 was approximately 3600 job incumbents. There are, as a consequence, small numbers of officers surveyed at any given grade level. The officer survey did, however, reveal some pattern differences in the kinds of tasks performed. Security police and legal officers, for example, performed more Uniform Code of Military Justice tasks than other kinds of officers (Bell, 1984). Also,

results indicated that junior officers spend a considerable amount of time on leadership, management, and communication tasks regardless of their AFSC. Reported time spent on these tasks ranged from 45% for pilot specialties to almost 80% for the logistics field. This finding is of particular interest in this effort, indicating that the taxonomy selected should include abilities most likely to account for performance in leadership, management, and communication tasks.

Another approach to task descriptions stems from McCormick's (1988, 1979) work with the Position Analysis Questionnaire (PAQ) which Cunningham (1988) expanded into the Occupation Analysis Inventory (OAI). Then, based on factor analysis of results of use of the OAI, Ballentine (Cunningham, Wimpee, & Ballentine, 1987) developed the General Work Inventory (GWI). These general purpose instruments, respectively, consist of 187, 617, and 268 items and are true task taxonomies. They elicit from job incumbents or SMEs ratings of the perceptual, job activity, interpersonal relationship, and physical elements which describe the task requirements of jobs. The approach represented by these job analysis techniques contrasts with the OSM approach because of the greater generality of the descriptive elements. Where the OSM would describe a job, for example, in terms of each of the planning tasks performed, the GWI describes at the level of asking respondents if the job requires Planning activity. The more generic nature of the descriptors, however, permits generalization across occupations which is essential to the application of a job component (or synthetic) validity approach to identifying ability requirements.

The PAQ, OAI, and GWI, since they are intended to describe jobs or specialties as opposed to tasks, provide broad coverage and are intended to support a variety of applications. The job elements of the PAQ and the OAI are arranged under six division headings, while the GWI job descriptive elements are organized under the 13 divisions shown in Table 1. Through principal components factor analyses, job dimensions for each of the methods have been identified. These job dimensions are used to determine factor scores for jobs to which the instruments are applied. These job dimension factor scores are required for achieving criterion levels of performance (Fleishman & Mumford, 1988).

Another behavioral description technique is Primoff's Job Element Method (JEM). Instead of using a written task analysis survey, as did the previously described techniques, the JEM approach generates the critical 'elements' of a job through SME conferences. The elements of a job consist of the knowledge, skills, abilities, and personal characteristics that each SME lists as important for successful performance. Each element is broken down into subelements, which are behavioral examples of an element. For example, the position of grocery store clerk might include elements such as accuracy, knowledge of stock, ability to be pleasant, and reliability. For the element accuracy, subelements may include accuracy in making change, accuracy in calculating cost, and accuracy in using the cash register. Each element and subelement is rated by the SMEs on four dimensions.

Table 1. General Work Inventory (GWI) Content

Sensory Activities (e. g., near vision, hearing)
Information Received or Used (e. g., written words, spoken words, readings from measurement or testing devices)
Information Related Activities (e. g., reading, difficult speaking, clerical operations)
Content of Information Used and Produced (e. g., mechanical information, electrical-electronic information, organizational management and administration information, engineering information)
Kinds of Thinking the Job Holder Engages In (e. g., expressing ideas, creating ideas, mathematical reasoning and problem solving, object problem solving and invention)
Physical Activities (e. g., working with fingers, tracking, strength)
Physical Activities Involving Tools, Equipment, etc. (e. g., small hand tools, measuring or testing devices, scientific or technical devices)
Work Performed With Tools, Equipment, Machines, or Devices (e. g., precision working, cutting by blade, abrading, cooking or preparing food, masoning, troweling, or casting)
Other Physical Work Activities (e. g., watching or monitoring tools or equipment, medical, health treating, caring)
Living Things Acted Upon (e. g., people, animals, plant life)
Materials Acted Upon (e. g., crude materials, finished materials, processes, substances)
Finished Products or Components Acted Upon (e. g., machines or mechanical equipment, structures, fabricated equipment)
Kinds of Contacts With People (e. g., managing or administering, supervising, evaluating, settling conflicts, communicating, teaching).
(Note: Each of the categories contain further definition and specific examples.)

From Work Inventory of Enlisted Specialties, AFPT 90-XXX-487, 1982.

The rating scales utilized in the JEM method are particularly useful for the purpose of selection. Ratings are collected for (a) whether or not a barely acceptable worker would have the element, (b) whether the element differentiates a superior worker, (c) whether there would be trouble if the element was not considered (whether the element affects performance), and (d) whether the element is practical to consider for selection (whether or not sufficient numbers of applicants would possess the characteristic described by the element). These ratings can be used in formulas developed to calculate various indices of value. The Item Index (IT) of an element describes the extent to which a subelement is important for establishing the content of an examination. The formula for IT is $SP + T$, where S = rating of importance in selecting superior workers, P = rating of practicality, and T = rating of trouble if not considered. Another item indice emphasizes the distinction between superior workers and others, as calculated by $S - B - P$ (B = extent to which barely acceptable workers possess the characteristic). The Total Value (TV) of an element describes the ability of that element to differentiate abilities of applicants for a job. The formula for TV is the combination of the two previous formulas, calculated for each SME, then summed across all SMEs. Other formulas describe the extent to which an element should be trained rather than assessed at the time of selection (Primoff & Eyde, 1988).

While the JEM approach incorporates abilities as well as task information, the elements are not true ability constructs. The abilities generated by this method may be more descriptive of procedural knowledge (ability to use a cash register) or personality (ability to be pleasant). Therefore this method is not sufficient in itself to determine the ability constructs which underlie the elements. In addition, the generation and rating of elements for each occupational specialty would be very costly and time consuming.

Of the behavioral description techniques reviewed in this study, Fine's Functional Job Analysis (FJA) method comes closest to identification of ability requirements. The FJA approach is based on task descriptions of over 4,000 jobs included in the Dictionary of Occupational Titles. Task verbs were sorted by job analysts into four functional categories: data, people, things, and worker instructions. These task related categories were related to three categories of ability: reasoning, mathematical, and language. Task verbs were further sorted within each ability category according to the complexity of the task; more complex tasks were assumed to require higher levels of the ability. This three-category taxonomy of abilities has six levels of reasoning ability, five levels of mathematical ability, and six levels of language ability. The most extensive application of this method is the occupational classification system of the US Employment Service. This method has also been used for test development, development of performance standards, development of training materials, job design, and job evaluation (Fine, 1988).

The FJA approach requires trained FJA analysts. The analyst(s) reviews existing job information and meets with SMEs. Task statements are then generated for the position of interest. Tasks are rated by the SMEs on frequency, importance, and criticality. Task verbs are then examined to

identify the level of reasoning, mathematical, and language ability required for the job.

This method could be applied to Air Force occupations by categorizing Air Force OSM task statement verbs to the levels within the three categories of ability. An advantage of the FJA approach is the extensive amount of information which has been collected using this method and the relation of task content to scores on the General Aptitude Test Battery (GATB). A disadvantage of this method is the limitations of the ability taxonomy. The three categories are quite broad. For example, the language ability scale includes speaking, reading, and writing abilities. The scales also include knowledge as well as ability. The lower levels of each ability are simplistic (common sense understanding, simple addition, follow simple oral instructions) and therefore not likely to be relevant to Air Force occupational tasks. Other abilities of interest are not included, such as perceptual and psychomotor skills. This approach would provide much more useful data if applied to a more refined taxonomy of abilities.

All of the behavior description techniques result in behavioral specifications of what workers actually do. Most of the behavior description techniques rely on written survey instruments. Instruments vary from detailed and specific (Air Force OSM task surveys) to general purpose instruments developed to be applied to any occupation (PAQ, OAI, GWI). The FJA and JEM methods rely on job analysts to collect task information.

The Air Force OSM task surveys are developed independently for each occupational specialty, which results in task statements which can be unique to a particular specialty. This prevents categorization of tasks across different specialties. An advantage to the general purpose instruments, such as the GWI, is that administration across various specialties would allow comparisons among the different specialties. A disadvantage to the general purpose instruments is that the detailed task information currently available through OMC task survey would not be utilized. Also, the general purpose instrument would have to be administered to samples of each occupational specialty, which would be a costly and time consuming endeavor.

The primary disadvantage of all of the behavior description methods is that they are lacking with regard to identification of abilities required to perform or learn the tasks successfully. An additional effort would be required in order to identify the prerequisite abilities for each task category of the written survey instruments. The JEM method is also lacking with regard to a taxonomy of abilities. Elements which are described as abilities are confounded with procedural knowledge (ability to use a cash register) and/or personal characteristics (ability to be pleasant). The FJA method does result in actual ability requirements. However, the taxonomy of abilities includes only three broad categories. A more refined and extensive taxonomy of abilities is desirable for application to Air Force occupational tasks.

2. The **Behavior Requirements Approach** places emphasis on the classification of behaviors that should be emitted or which are assumed to be required for achieving criterion levels of performance (Fleishman & Mumford, 1988). The human performer is assumed to be in possession of a large number of processes that serve to intervene between the stimulus and response events. These intervening processes are in a real sense constructs to account for human task behavior. Methods based on this approach provide classification descriptors that more closely resemble a taxonomy of abilities than do behavioral descriptive methods. Behavioral requirements methods have most generally been applied to the problem of human learning, two of the more representative of them being employed by Gagne (1977) and Miller (1973).

Gagne directed his efforts at the field of human learning, contending learning consists not only of the content or subject matter but also the process by which existing knowledge and skills are modified. Gagne proposes five categories of learning processes: (a) intellectual skills, (b) cognitive strategy, (c) verbal information, (d) attitude, and (e) motor skills. Intellectual skills involve discriminations, concepts, rules, and their elaboration. Cognitive strategies are the internally organized skills governing the individual's behavior in learning, remembering, and thinking. Verbal information involves learning facts, principles, and generalizations required in the learning of a particular content area. Attitude is another type of learning process and is generally least affected by training. Motor skills are the capabilities mediating organized motor performance that can be improved through practice. Tasks can be classified according to this scheme, thus providing the trainer with the basis of a training program. Application requires highly trained specialists to analyze and classify each task according to the taxonomy of intervening processes.

Miller (Fleishman & Quaintance, 1984) classifies tasks according to 25 task functions involved in a generalized information-processing system. Some of the 25 elements are Message, Input, Identify, Short Term Memory, Interpret, and Goal Image. For example, Message is "A collection of symbols sent as a meaningful statement" (p. 438); Input is "Selecting what to pay attention to next" (p. 439); Identify is "What is it and what is its name" (p. 444); Short Term Memory is "Holding something temporarily" (p. 450); Interpret is "What does it mean" (p. 445); and Goal Image is "A picture of a task well done" (p. 459). A complete description along with detailed instructions for the use of each function are in Appendix A in Fleishman and Quaintance (1984). Tasks, according to Miller, can be classified according to these functions and such classification can be useful for selection and training.

In summary, Behavioral Requirements approaches identify the intervening behavioral processes required for task performance. They do not provide for identification of abilities, although the information provided could serve as a basis for such identification.

3. The **Task Characteristics Approach** differs from the other taxonomic approaches in that the emphasis is on characteristics of the task itself. Behavior description taxonomies describe tasks in terms of worker activities, such as 'repair electronic ignitions.' Behavior requirement taxonomies categorize tasks in terms of worker functions, and the ability requirement taxonomies describe worker abilities. In contrast, the task characteristic approach classifies tasks in terms of characteristics of the task which impact on behavior and performance. Basic to the identification of task characteristics is the definition of the concept task, described by Hackman (1970) as

"A task is assigned to a person (or group) by an external agent or is self-generated, and consists of a stimulus complex and a set of instructions which specify what is to be done vis-a-vis the stimuli. The instructions indicate what operations are to be performed by the performers with respect to the stimuli and/or what goal is to be achieved (p. 210)."

According to this definition there are three aspects of tasks to be considered: (a) type of stimulus materials, (b) the instructions about the goals, and (c) the instructions about procedures. A task must have stimulus materials. For example, the instruction 'write' would not be a task, but the instruction 'write a summary which describes your job duties' would be a task. Tasks should have instructions, either regarding procedures and/or goals.

Task characteristics are generated on the basis of attributes of the stimulus and instructions. For example, three task characteristics have been identified as contributing to the overall stress of a task. One is time sequencing, either of stimulus (high versus low rate of presentation of stimulus) or instructions (high versus low rate of response, strict time deadlines, or repetitive response). Another characteristic affecting task stress is complexity-interpretability, either of stimulus (high versus low amount of complexity or ambiguity) or instructions (high versus low amount of complexity or ambiguity). The third characteristic affecting task stress is threat or punishment, inherent in the stimulus (run across a rope bridge 200 feet above the ground) procedural instructions (drive a car at high speed), or goal instructions (blow up a balloon until it breaks) Hackman, 1970. This method of assessing task attributes may group tasks that are very different with regard to activities, worker functions, or worker abilities, as in the examples of stressful tasks.

A taxonomy of task characteristics has been developed and consists of 21 individual scales (Fleishman & Quaintance, 1984). Examples of task characteristics include (a) dependency of procedural steps, (b) number of procedural steps, (c) response rates, (d) procedural complexity, and (e) feedback. Task characteristics have been shown to have important implications for the prediction of task learning and retention (Rose et al., 1984; Wheaton &

Mirabella, 1972). The task characteristic approach may have great potential for the identification of ability requirements, especially if particular task characteristics are identified or generated to reflect certain ability requirements. However, there has been no reported efforts of this nature, nor is there a model for such an application. Indeed, it would be a massive effort to apply this approach to Air Force task data; all tasks would have to be assessed using SMEs to identify all task characteristics of each task. Even then, the link to ability requirements would have to be established. Thus, this approach was dismissed from further consideration.

4. The Abilities Requirements Approach describes, contrasts, and compares tasks in terms of the abilities they require of the operator. Abilities as defined by Fleishman and Mumford (1988) and McCormick (1979) are the relatively enduring attributes of the individual performing the tasks, and certain tasks are hypothesized as requiring certain ability profiles if performance is to be maximized. Fleishman and his associates (Fleishman & Mumford, 1988; Fleishman & Quaintance, 1984; among others) have developed a general purpose method of assessing the ability requirements of task performance through application of their Ability Requirements Scales (ARS). The abilities included in the ARS constitute the taxonomy. Pre-existence of tasks describing the job is assumed, since no methods for identifying the tasks of the job being analyzed by the Ability Requirement Scales are provided. Basing the analysis on tasks is, however, recommended.

Fleishman's approach differs from all others in very significant ways. Beginning with the premise that different tasks require different ability profiles, he defined and tested various constructs of human abilities in the cognitive (taken from the factor analytic literature, principally that of Guilford and French), perceptual, psychomotor, and physical domains. The result is a scheme for classifying tasks according to a taxonomy of 37 to 52 human attributes, depending upon which version of his instrument is used. The difference in his approach lies in that he derived the attributes from the factor analytic literature, identified or developed tests of the attributes, and then tested the application of the attributes in the description of tasks in the laboratory as well as in actual practice. Only after the extensive testing to establish the construct validity of the ability factors did Fleishman then begin to analyze tasks in terms of the abilities required to perform them. Fleishman, thus, provides a set of defined sensory, perceptual, cognitive, psychomotor, and physical ability factors. In contrast, others (Cunningham, 1988; Fine, 1988, and McCormick et al., 1977; for example) begin with a description of jobs or tasks, and then by a variety of methods identify the abilities required. They have no predefined set of abilities; the job analyst determines the list of abilities or ability measures to be employed.

Fleishman's taxonomy of cognitive abilities consists of 19 human attributes, the final number having been derived empirically from studies comparing use of a small number of generally defined abilities with the larger number of more specifically defined abilities. The cognitive abilities consist of such attributes as Oral Comprehension, the ability to understand spoken

English words and sentences; Written Comprehension, the ability to use English words or sentences in speaking so others will understand; Fluency of Ideas, the ability to produce a number of ideas about a given topic; Deductive Reasoning, the ability to apply general rules to specific problems to come up with logical answers; Inductive Reasoning, the ability to combine separate pieces of information to form general rules or conclusions; and Spatial Orientation, the ability to tell where one is in relation to the location of some object or to tell where the object is in relation to oneself.

In addition to the 19 cognitive abilities, there are sensory, perceptual, psychomotor, and physical abilities included in the ARS. A complete description of Fleishman's human ability dimensions is shown in Fleishman and Quaintance (1984) and Fleishman and Mumford (1988); Table 2 lists the 52 abilities. Table 3 provides an example of one of the ARS scales and definitions.

While the ARS has used a wide variety of applications, one has special relevance for Air Force use. Landy (1988) stated that the first obstacle in using any ability taxonomy is conceptual in that the typical SME is not familiar with many of the concepts or labels that are used. To increase SME familiarity, Landy provided instruction on each of the ability areas by using salient examples of each drawn from the occupations from which the SME came. If, for example, the position to be studied was fire captain, examples were developed for each ability from the job of fire captains. These examples were incorporated into the definition of each ability. In addition, Landy devised a different scale for the ARS than is provided. SME estimated the importance of each ability using a 100-point allocation method. SME were instructed that for each area rated, 100 points were to be allocated among the set of abilities required of the area.

Landy's approach to SME training would certainly be expected to produce greater interrater reliability, because of the raters' greater familiarity with the ability definitions and their relationships to work requirements. For this reason it has a great deal of merit. The disadvantage, however, is the cost in developing the job related ability definitions and the training of the raters.

The scale Landy employed has to be viewed from the perspective of what information about ability requirements is needed and the design of the rating process. If knowledge of the relative requirement of each of the abilities of interest for each task or job is the objective, his scale is appropriate. In terms of the rating process, his approach requires the rating of each task or job on the relative requirement of each of the total list of abilities. However, it eliminates any form of absolute rating and inhibits use of an equal-interval scale.

Although they did not use the complete ARS, Peterson and Bownas (1982) have applied an abilities requirement taxonomic approach that has potential

Table 2. Ability Requirement Scales

Cognitive

Oral Comprehension
Oral Expression
Fluency of Ideas
Memorization
Mechanical Reasoning
Deductive Reasoning
Speed of Closure
Spatial Orientation
Category Flexibility
Perceptual Speed

Written Comprehension
Written Expression
Originality
Problem Sensitivity
Number Facility
Inductive Reasoning
Flexibility of Closure
Visualization
Information Ordering

Perceptual

Multilimb Coordination
Rate Control
Control Precision

Reaction Time
Time Sharing
Response Orientation

Psychomotor

Arm-Hand Steadiness
Finger Dexterity
Speed of Limb Movement

Manual Dexterity
Wrist-Finger Speed
Selective Attention

Physical

Static Strength
Trunk Strength
Gross Body Coordination
Dynamic Flexibility
Stamina

Explosive Strength
Dynamic Strength
Gross Body Equilibrium
Extent Flexibility

Sensory

Near Vision
Visual Color Discrimination
Peripheral Vision
Glare Sensitivity
Auditory Attention
Speech Hearing

Far Vision
Night Vision
Depth Perception
General Hearing
Sound Localization
Speech Clarity

Table 3. An Example of Ability Requirement Scales Definition

VERBAL COMPREHENSION

This is the ability to understand English words and sentences.

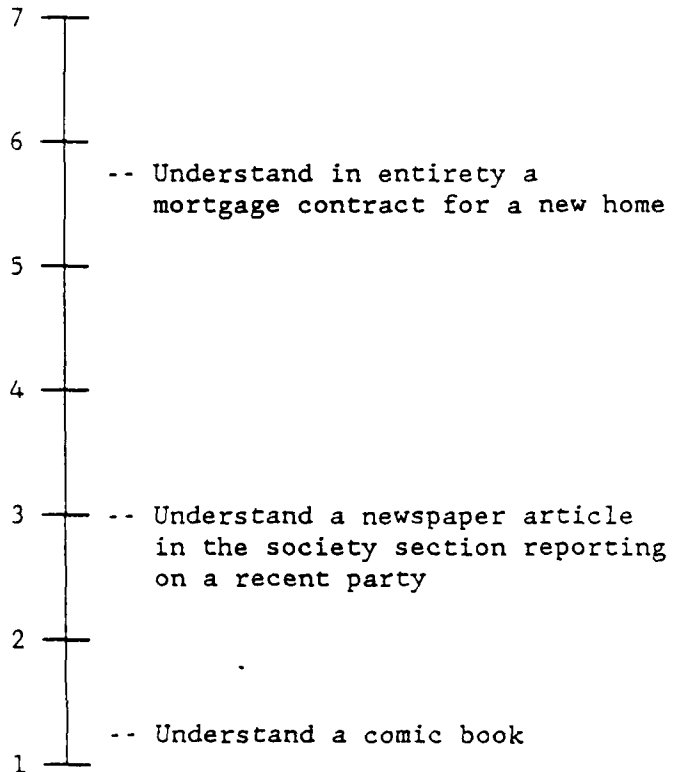
How Verbal Comprehension Is Different From Other Abilities

Understand spoken or written
English words and sentences

vs.

Verbal Expression: Speak
or write English words or
sentences so others will
understand.

Requires understanding of complex,
detailed information which contains
unusual words and phrases
and involves fine distinctions
in meaning among words.



Requires a basic knowledge of
language necessary to understand
simple communications.

Theologus, 1970

relevance for Air Force application. They employed the ARS psychomotor and physical ability scales, but substituted a different set of cognitive abilities for the ARS cognitive factors. The cognitive factors consisted of Dunnette's (1976) 10 factors plus Verbal Closure and Visual Memory abilities (see Table 4 for a listing of the 12 cognitive abilities). In addition, they also included 15 personality dimensions and six vocational preference dimensions. The particular relevance of this approach lies in the reduction of the number of cognitive abilities to be rated.

Table 4. Cognitive Abilities

Cognitive Abilities

Flexibility and Speed of Closure
Fluency
Inductive Reasoning
Associative (Rote) Memory
Span Memory
Number Facility
Perceptual Speed
Logical Reasoning
Spatial Orientation and Visualization
Verbal Comprehension
Verbal Closure
Visual Memory

The ARS scales have a particular advantage in that they have been used successfully in a number of military studies. In a study performed for the Army Research Institute (ARS) ability constructs were included in an initial list of knowledge, skills, and abilities (KSAs) hypothesized to be related to leadership skill. The final taxonomy of KSAs which were related to leadership included many of the ARS ability constructs (Mumford, Yarkin-Levin, Korotkin, Wallis, & Marshall-Mies, 1986). A study performed for the Navy also utilized ARS constructs as one of four methods to identify variables most likely to predict job performance. It was found that SMEs had no trouble in comprehending ARS ability definitions and were able to provide reliable judgments of ability requirements, although the rating scale provided by the ARS was not employed. Instead, a 5-point scale to assess consequences on performance if workers did not possess the ability at a desired level was used. The ARS method and a policy-capturing method were found to be most effective in identifying optimal predictors. Results indicated that the predictive validity and utility of different types of predictors (cognitive ability, psychomotor, information processing, previous training and experience) will vary among occupational specialties. The study concluded that while cognitive ability predictors would be likely to predict performance to some extent for all specialties, other predictors may be more appropriate for some occupational fields (Murphy, 1988).

The ARS ability constructs have also been utilized in previous Air Force studies. Perceptual and psychomotor abilities have been related to pilot tasks, and measures were developed for a computer-administered experimental battery (Kantor & Bordelon, 1985). A separate Air Force study investigated taxonomies of perceptual and psychomotor abilities in order to identify the optimum set of constructs for application to Air Force occupational tasks. The resulting list of 13 perceptual and psychomotor constructs included seven ARS constructs and six from other sources (Siegel, Federman, & Welsand, 1980). Ratings were collected for each of the abilities for 35 enlisted Air Force occupations. Results indicated that the methodology utilized was successful and that perceptual and psychomotor ability profiles varied for different career fields.

An Army study utilized ARS scales in a predictive validity study. Aviation ability tests were developed on the basis of ability requirements determined through application of the ARS method. Experienced helicopter instructor pilots identified the critical tasks of helicopter pilot specialties. The tasks were rated by the SMEs with regard to the ARS abilities required to perform each task. Twelve abilities were chosen for test development. The experimental battery was administered to entrants of helicopter pilot training (N = 563). Follow up performance data was subsequently collected. Results indicated that the experimental battery accounted for 27% of the variance in performance measures (McAnulty, 1988).

Another method which fits into the ability requirements approach was developed by Lopez (1988). His method, the Threshold Traits Analysis System (TTAS) analyzes jobs through observation, using a taxonomy of job requirements to describe each job. Lopez proceeded to develop his TTAS to specify the human traits required to perform different classes of job functions. His TTAS consists of five areas: physical, mental, learned, motivational, and social. The first three refer to abilities. Each consists of a number of traits, each of which is operationally defined. For example, mental ability is divided into vigilance, attention and information processing. Information processing is in turn described as consisting of memory, comprehension, problem-solving, and creativity. Memory is operationally defined in terms of whether the worker can "Retain and recall ideas;" comprehension is operationally defined as "Understand spoken and written ideas" (p. 882). In all, the TTAS provides a list of 16 generally defined abilities which are applied at the job level through the observation of workers.

Special Purpose Taxonomies

Special purpose taxonomies designed for management positions were also reviewed, because of their potential value for describing officer positions. Since about 1950, there have been a number of factor analytic studies directed at identifying common work dimensions with the hope of categorizing managerial jobs according to similarities in their factorial structure. Summarizing this work, Fleishman and Quaintance (1984) indicate that managerial jobs do have

common work dimensions that may be used as the basis for a classificatory system.

Generally, about a dozen dimensions of managerial work have been identified and, as Landy and Trumbo (1980) indicate, these dimensions have been replicated many times with different instruments, theoretical approaches, and populations, and they are intercorrelated. Two overall factors, Consideration and Structure, account for over 80% of the variance of leadership behavior. A considerable body of the research stems from Hemphill's (1959) work that identified 10 factors. In later research using a modified version of Hemphill's instrument, Tornow and Pinto (1976) identified 13 factors. These factors were derived from data collected from upper, middle, and lower level management personnel on a 204-item Management Position Description Questionnaire (MPDQ).

More recently, Page (1988) updated the MPDQ. It is a Behavioral Description approach, consisting of 274 job behaviors (including 31 knowledge, skill, or ability, 16 general information, and five decision making elements) to which respondents use a 5-point scale to indicate the significance of the behaviors to their positions. Factor analysis of data obtained from administration of the MPDQ in the business sector revealed the following eight management work factors: decision making; planning and organizing; administering; controlling; consulting and innovating; coordinating; representing; and monitoring business indicators. In addition, nine management performance factors were identified: managing work; business planning; problem solving and decision making; communication; customer and public relations; human resource development; human resource management; organizational support and interface; and job knowledge. Reliability of the instrument reported for seven studies ranged from .79 to .93. The MPDQ is a general purpose instrument that describes management positions at a generic level, thus generalization across positions may be made.

Another approach to the study of management jobs is through use of the Professional and Managerial Position Questionnaire (PMPQ) (J. L. Mitchell, personal communication on January 20, 1989). The PMPQ consists of a set of descriptors applicable across management jobs. It bears a close resemblance to the PAQ, which is not surprising since Mitchell was a student of McCormick. In the PMPQ, respondents indicate, first, what part of their job (10-point scale with 0 denoting does not apply) a given descriptor represents; and, second, for those descriptors involved in their jobs, respondents rate the level of complexity (on a similar 10-point scale). The Navy Occupational Development and Analysis Center is currently administering a tailored version of the PMPQ to a sample of 10,000 Navy officers, grades W-2 through O-6, in all ratings in a study of the similarity of Navy Officer jobs (J. Treckel, personal communication on December 5, 1988).

It is important to note that none of the efforts to study management jobs, including the OSM surveys, had as its objective the identification of ability

requirements. Methods used to describe jobs, however, are highly similar to approaches for "non-management" jobs discussed in earlier sections. Classification of, as opposed to initial selection for, the jobs studied was the objective. There is no management ability counterpart of a taxonomy similar to the ARS.

That there is no such taxonomy is not surprising. The scope of the various cognitive, perceptual, psychomotor, and physical abilities (as well as the personality and vocational preference dimensions employed by Peterson and Bownas) leaves few, if any, abilities measured in traditional ways untapped. Inspection of the factors reported by Page (1988) suggests that the abilities so far defined are applicable to management jobs. However, they may be applicable and yet not sufficient in themselves to identify all abilities which underlie leadership management and communication tasks. Therefore, taxonomies purported to measure managerial positions are scrutinized here for their relevance in a taxonomy for Air Force occupational specialties. The skills required for leadership, management, and communication tasks may be more complex and higher level than those included in other taxonomies, such as ARS. It will be a challenge to identify and define the best set.

The problem of identifying ability requirements for officer jobs seems to lie not in a lack of a taxonomy of abilities or of measures of various sensory, cognitive, perceptual, psychomotor, or physical abilities but in the manner in which the jobs are to be described and in the inferential leap to link these descriptions to the requisite abilities. Even here, the problem is not crucial, for there are OSM data for about half of the officer specialties and at least six descriptive methods that can be used to describe the remainder.

Five of the seven methods discussed in the preceding sections (FJA, JEM, PAQ, OAI, GWI) are descriptive methods, and the MDPQ and the PMPQ are descriptive methods as well. Of these techniques, only the GWI can be categorically identified as not having been employed to study management or officer-equivalent jobs, although it has been used to survey enlisted incumbents. Inspection of the content of the GWI, however, suggests that it includes job elements which are entirely appropriate for describing officer jobs. For instance, consider these examples of groupings of GWI descriptive elements in Table 1:

Information Received or Used (e. g., written words, spoken words, readings from measurement or testing devices); Information Related Activities (e. g., reading, difficult speaking, clerical operations); Content of Information Used and Produced (e. g., mechanical information, electrical-electronic information, organizational management and administration information, engineering information); Kinds of Thinking the Job Holder Engages In (e. g., expressing ideas, creating ideas, mathematical reasoning and problem solving, object problem solving and invention); Kinds of Contacts With People (e. g.,

managing or administering, supervising, evaluating, settling conflicts, communicating, teaching).

Descriptive elements of these kinds seem as applicable to officer specialties as they are to enlisted specialties and equally as easily linked to ability requirements.

In summary, the job analysis methods designed to study professional management jobs provide little, if any, more descriptive information relevant for describing Air Force jobs than do the general purpose descriptive methods. Use of one of these latter methods, the GWI, should provide sufficient information for describing specialties for which OSM task data are unavailable.

Rating Scales: Measurement and Procedural Issues

The rating scales used to determine ability requirements vary among the different methods. The methods reviewed differed with respect to (a) who uses the rating scales (job analysts, incumbents, supervisors), (b) how the scales are administered (in SME conferences or by mail), (c) the nature of the scales (dichotomous, ordinal, absolute), and (d) the relevance of the scale to selection. In this section, the rating scales will be discussed with regard to the differentiating factors. The first two characteristics are important in their effect on the cost of the procedure. The use of specially trained job analysts would greatly increase the cost, as would the use of SME conferences. The last two characteristics influence the effectiveness of the information that is derived.

FJA (Fine, 1988) and JEM (Primoff & Eyde, 1988) employ trained analysts to facilitate the process by which SMEs identify the ability requirements. The OAI (Cunningham, 1988) employs trained analysts and trained SMEs. The PAQ (McCormick et al., 1977), GWI (Ballentine & Cunningham, 1981) and ARS scales (Fleishman, 1988) can be used by job analysts, supervisors, or incumbents (if incumbents are fairly well educated). It is not certain which group is the 'best' source of information. Previous studies which examined differences between supervisors, incumbents, and analysts produced conflicting results, perhaps due to differences in the method used (Cornelius, 1988). It was concluded that incumbents may tend to inflate ratings relative to ratings by supervisors, and that both incumbents and supervisors may tend to inflate ratings relative to analysts. This tendency was found when the job elements rated were high in social desirability, and when scale points were not sufficiently described.

The amount of time and effort involved in the administration of rating scales differed considerably among the methods. The PAQ, OAI, and GWI are all administered by mail with a minimum of effort. The FJA and JEM methods rely on SME conferences facilitated by trained analysts. The TTA method requires

direct observation of incumbents by trained analysts. The ARS scales have been applied by mail and through SME conferences. Fleishman (1988) and Landy (1988) recommend the use of SME conferences with ARS scales. Other studies have reported success using ARS scales by mail (Mumford et al., 1986; Murphy, 1988; Siegel et al., 1980). While it is expected that SME training that occurs in the conference situation would increase interrater reliability, some studies have found that the questionnaire approach yielded information equivalent to that collected by SME conference, when the survey included a large number of respondents (Levine & Sistrunk, 1988). Administration by mail may be suitable for a pilot study, in order to determine whether a need exists for SME training in a conference situation.

Rating scales used by the different methods differ widely in terms of what is actually being rated, and the quantitative nature of the scale. A primary difference is whether the rating scale is applied to describe tasks, as in the behavior description methods, or ability requirements. Rating scales for task characteristics are fairly similar, tapping information about whether or not the task is performed, the frequency or time spent on the task, and the importance of the task. Additional task ratings may be collected for the determination of training requirements or task difficulty.

The Air Force OSM task ratings typically provide information regarding the percent of incumbents performing the task, the percent of time spent on the task of those who performed the task, and the percent of time spent across all incumbents. Additional ratings are also collected for specific purposes, such as the determination of training curriculum. The PAQ uses five-point scales to describe the extent of use, importance to job, amount of time, and possibility of occurrence. The OAI uses five-point scales that describe the significance, applicability, and extent (time spent) of each item. The GWI uses a nine-point scale which describes the extent to which the item is part of the job. The FJA method may collect ratings of frequency, importance, and criticality, if the user wishes the information. However, Fine (1988) stated that the usefulness of the scales is doubtful and that the critical issue is whether or not the task is performed at all.

For application to the purpose of selection, the main objective of task data is that the tasks most representative (% incumbents performing) and important for successful performance of the job are identified. Also desirable is information that describes whether tasks are to be trained, so that determination of the ability requirements will either be based on the ability to perform the task without training, or the ability to learn the task. This issue should be addressed either at the point of selecting tasks to be rated, or within the instructions for the SMEs.

Rating scales used to determine ability requirements vary with regard to what is being rated and the type of scale utilized. The JEM rating scales can be applied to abilities or tasks, since the job elements identified through JEM can be in the form of knowledge, ability, or personal characteristic. The four

JEM scales are ordinal, using three-point scales to determine (a) the extent to which barely acceptable workers possess the characteristic (all, some, almost none), (b) the extent to which the characteristic is particular to superior workers (very important, valuable, or does not differentiate), (c) whether trouble is likely if the characteristic is not considered (much trouble, some trouble, or safe to ignore) and d) whether the characteristic is practical to consider for selection (by demanding this characteristic we can fill all openings, some openings, or almost no openings). In Lopez's (1988) TTA method elements are rated 1 or 0 to indicate importance to task performance, and 1-3 to indicate uniqueness and trait level.

Fleishman's ARS method uses unique seven-point scales which describe the levels of each ability to be considered by raters. Each level of an ability is benchmarked by behavioral examples, in order to produce an absolute scale of ability requirements. The behavioral examples were developed using a behaviorally-anchored rating scale methodology. These scales should allow comparisons of the level of ability required across jobs, without the inflation that is typical of ordinal scales. However, there is doubt as to the absolute nature of the scales. McNulty and Jones (1984) reported that the raters used the scales as if they were ordinal, rather than absolute.

Landy (1988) suggests a different rating format to be used with the ARS scales. He recommended that the standard behavioral examples for the levels of abilities be supplemented by behavioral examples taken from job content. Thus, the Fleishman description for a high level of reading comprehension ability (read and comprehend a home mortgage) would be replaced or supplemented by a more job-related example. In addition, Landy utilized a point allocation system to identify the abilities required for task clusters. The point allocation system requires the rater to distribute 100 points across the abilities so that they represent the importance to a particular task. Points for each ability are averaged across task clusters, resulting in percentages of ability requirements for the job as a whole. This enables the identification of characteristics which may be particularly important to the job. For example, for the position patrol officer raters reported that verbal skills accounted for 30% of the ability requirements, and that cognitive skills in general accounted for 81% of the ability requirements.

The rating scales developed for JEM have particular relevance for the purpose of selection. The four rating scales are utilized in formulas developed specifically to evaluate the importance of the characteristic to selection. The Item Index (IT) of an element describes the extent to which a subelement is important for establishing the content of an examination. The formula for IT is $SP + T$, where S = rating of importance in selecting superior workers, P = rating of practicality, and T = rating of trouble if not considered. Another item indice is that which emphasizes the distinction between superior workers and others, as calculated by $S - B - P$ (B = extent to which barely acceptable workers possess the characteristic). The Total Value (TV) of an element describes the ability of that element to differentiate abilities of applicants for a job. The formula for TV is the combination of

the two previous formulas, calculated for each SME, then summed across all SMEs. Other formulas describe the extent to which an element should be trained rather than assessed at the time of selection, whether the element should be used for screenout (minimum requirement), and whether the element should be used for ranking applicants. Tests are then developed to assess specific elements identified as useful for selection (Primoff & Eyde, 1988).

The JEM approach uses a slightly different scale when test development is not feasible, in order to relate the test to the job elements. The procedure results in a J-coefficient, based on a synthetic validity approach. The rating scale applied to each element consists of a single three point scale: 2 points if most of the superior workers have the ability and most of the unsatisfactory workers do not have the ability; 1 point if workers with the ability will be better than workers without it, but the ability is not absolutely necessary, and 0 points if there is no relation between the ability and successful performance. The summed points for a particular element is considered its weight of the element in a criterion of job success. Standard worksheets are then used to estimate the expected validity coefficient(s).

Because the JEM rating scales were developed for the enhancement of selection decisions and take into account issues such as training after selection and practicality for selection, they merit serious consideration for application to Air Force occupational tasks. A factor to consider is the extent to which the rating time is increased as a result of using all four scales. The single three point scale may be more appropriate in order to reduce the demands on the raters. Another factor is the impact of the special scales on interrater reliability. Raters may agree on whether an ability is required, but may not be in agreement as to the level of ability required or to what extent applicants possess the ability.

An optimal approach would be the application of JEM scales with ARS scales, modified by job-related behavioral examples and SME training as recommended by Landy (1988). The ARS scales provide explicit descriptions of each ability along with behavioral examples, while the JEM scales describe the relevance of that ability to selection. The ARS scales are based on ability constructs derived from rigorous discriminant validity studies, and tests of each ability construct are available. The JEM scales may be used to calculate indices describing the relevance of each ability to selection and training. However, this modified approach would require additional effort and expense, in order to produce the behavioral examples, train the SME raters on the process, and have the raters rate on four dimensions.

Internal Validity: Interrater Reliability and Taxonomic Adequacy

There are two important indices of internal validity: reliability and extent or exhaustiveness of coverage. Generally, interrater agreement has been reported as high. With two trained analysts, Cunningham (1988) reports the

median interrater agreement from a study of 215 jobs to be .82 for the OAI. For the GWI which was administered through a mail survey, Cunningham, Wimpee, and Ballentine (1987) found a mean rate-rater agreement of .62 on individual items and .66 for item profiles. For the PAQ McCormick and Jeannert (1988) report that, using two analysts, the "... average reliability coefficients have typically been in the .80s..." (p. 830). Rate-rater reliability studies have produced coefficients as high as .80. Using two raters who judged 33 traits on 100 jobs, Lopez found a median split half reliability coefficient corrected by the Spearman-Brown Prophecy Formula of .86. Agreement among two samples of six raters each on "job criterion values" (or the sum of the job analysis ratings) was .92. Primoff and Eyde (1988) reported interrater reliability for job elements across 50 studies was .86, after correction using the Spearman-Brown formula. The FJA typically employs one analyst, although Fine (1988) reports "... high reliability was the result of precise descriptions (of tasks)" (p. 1023) in a developmental study involving 10 raters judging 50 jobs.

The most extensive study of reliability is reported for the use of the Ability Requirements Scales developed by Fleishman and his associates (Fleishman & Mumford, 1988; Fleishman & Quaintance, 1984). A major concern in the development of the methodology was the achievement of high reliability and considerable experimental work was accomplished, including scale and factor redefinitions. In practice, at the job level, 19 Ability Requirement Scales were used to examine 15 benchmark jobs. Interrater reliability coefficients "... were typically in the low .90 range and fell below .80 in only a single occupational field ..." (Fleishman & Mumford, 1988, p. 928). Substantial agreement was also found between different groups of raters. Agreement between incumbents and supervisors, except for one job, was above .75. Numerous studies employing the Ability Rating Scales are reported in Fleishman and Quaintance (1984) in which agreement among different groups of raters ranged from .50 to .87, with the majority of the coefficients being above .70. They conclude: "Interrater agreement with current versions of these scales, used to describe jobs and tasks in a wide variety of industrial, governmental, and military settings, tend to be in the .80s and .90s" (p. 330). It should be noted that Fleishman recommends using at least 20 raters, although he reported instances where acceptably high interrater agreement was found for fewer raters. His recommendation on number of raters needed is consistent with the present practices in the OSM methodology in which at least 30 raters are sought to assure interrater agreement at the .90 level.

Methodologies differ widely with regard to the set of abilities considered for the establishment of job requirements. Behavior description methods provide little if any information about ability requirements. USAF/OMC task surveys are strictly task statements. A separate taxonomy of abilities must be used in order to establish ability requirements. The PAQ and OAI have been related to nine measures of cognitive ability within the General Aptitude Test Battery (GATB). As yet, the GWI has not been referenced to any ability measures. The FJA method relates task statements to three broad categories of ability: reasoning, mathematical, and language ability. The JEM may produce any number of abilities, with definitions specific to the job being described.

The JEM method has also been related to GATB scores and tests developed to measure specific elements (Primoff & Eyde, 1988).

Of the ability taxonomies reviewed in this study, the taxonomy developed by Fleishman has the broadest range of abilities and demonstrates the highest degree of construct validity. The ARS taxonomy in its longest form consists of 52 abilities, 19 of which are cognitive abilities. No task descriptive items are included. Initial development of the cognitive abilities began with a study of the work of Guilford and others about cognitive abilities (Fleishman & Mumford, 1988; Fleishman & Quaintance, 1984) to identify the abilities required for human task performance. Tests for each of these abilities having construct validity for measuring an ability were identified. The criterion for construct validity was evidence of such validity in at least 10 independent research studies.

The ARS taxonomy is not without some disadvantages. The ability constructs included in the ARS taxonomy are fundamental; more complex abilities which are better described in terms of the task performed (ability to plan, ability to manage time, etc.) are not included. It is assumed that higher level job skills are driven by the basic abilities contained in the ARS taxonomy. This is an advantage if the assumption is true, because tests of the ability constructs can be used off the shelf without tailoring to the job. However, if a test of an ability were to be developed, it is recommended that the tasks related to the ability be reviewed. Items can be constructed which measure the ability of interest, such as reasoning, using problems based on tasks which required the ability.

The ARS taxonomy can be modified to be more useful to Air Force occupations. Some of ability constructs can be deleted, such as the physical abilities. The set of cognitive abilities can be replaced with a smaller set, such as those described by Peterson and Bownas (1982) or Mumford et al. (1986). SMEs could be consulted in order to reduce the set of constructs. The perceptual and psychomotor constructs should consist of those identified by Siegel et al. (1980) as relevant for Air Force occupations. Some ability constructs should be added, particularly those which have been related to successful performance of leadership, management, and communication tasks. Communication and interpersonal skill constructs which should be considered for inclusion are oral presentation ability, influence (ability to persuade others), and interpersonal problem solving ability. Rating scales and behavioral examples would have to be developed for additional ability constructs.

External Validity: Job Component and Criterion-related Validity

External validity pertains to how well the classificatory system achieves the objectives for which it is intended. The classificatory system should differentiate among tasks or jobs in terms of their ability requirements and

these requirements should be valid--that is, the requirements as stated should be possessed by job candidates selected for the job in order to provide reasonable assurance that these candidates will become satisfactory performers. In this regard, McCormick (1979) states, "Validity ... is not a clear-cut, unitary concept; rather, there are different varieties of validity; it is usually not feasible to refer to any given job requirement as valid or not valid. Instead one needs to think in terms of the degree of validity or, perhaps more practically, to view validity in terms of the probabilities that those fulfilling a stated requirement will become satisfactory employees. Even the term satisfactory is fairly slippery, since in the case of most jobs the incumbents vary in degree of satisfactoriness" (pp. 245-246).

According to McCormick, in circumstances where the prediction deals essentially with inexperienced individuals to be trained from scratch on a given job, the primary interest is in predicting the suitability of individuals for learning and adapting to a job. This prediction is largely formed on the basis of abilities or other attributes. Any tests that are used to measure candidates on the attribute requirements of jobs preferably should have substantial construct validity. That is, they should measure with "reasonable fidelity" the basic attributes or constructs which they are intended to measure. Evidence that a construct measured by a test is relevant to a given job needs to be supported by criterion-related validity "... or can be inferred from a sound job analysis" (McCormick, 1979, p. 246). He indicates that both the Uniform Guidelines (1978) and good professional practice focus attention on the potential utility of sound job-related data as the most justifiable basis for establishing job requirements.

There has been considerable effort to establish job component (or synthetic) validity, although the specific results of these efforts do not appear to be applicable to the Air Force selection process. There are, however, implications for the use of job component validity for identifying ability requirements for AFS, if such validity is based on Air Force research. Typical of some of the efforts to establish component validity are those reported by McCormick (1979) and Cunningham (1988). McCormick employed multiple regression analysis, using job dimension scores from the PAQ as predictors of scores on each of the nine tests on the GATB. This approach required a sample of jobs for which GATB scores were available for job incumbents and for which there were data from PAQ analyses of the jobs. The approach is predicated on the assumptions that jobs differ with respect to which a given attribute is required for successful performance and that workers tend to gravitate to jobs which are commensurate with their abilities. On these assumptions, then for a given attribute the differences among mean test scores of workers in different jobs would reflect in a gross way the varying job requirements for that attribute for the jobs being studied. McCormick (1979) reports multiple correlations of PAQ factor and GATB test scores ranging from .30 (for manual dexterity) to .83 (for verbal) for the nine GATB tests. The median multiple correlation was .73. McCormick et al. (1977) also rated PAQ job dimension scores to 19 commercially-available tests that were considered to match the GATB subtests. Slightly lower multiple correlation coefficients were found.

Cunningham (1988) reports similar results from use of the OAI. In addition to OAI ratings for jobs, he obtained ratings of the 617 OAI elements on 103 attributes as McCormick (McCormick et al., 1977) had previously done for the PAQ, thus giving two kinds of OAI job ratings: (a) in terms of scores on factors representing different kinds of work activities and conditions; and (b) in terms of estimated requirements for various human attributes for which there are tests. Using GATB test scores, Cunningham compared clusters of jobs with similar OAI factor score profiles in terms of mean test and inventory scores for relevant samples of job incumbents. Analysis of variance results showed statistically significant discriminability among clusters for 68 of the 92 measures employed. Also, multiple correlations between OAI factor scores for a sample of jobs and the GATB scores produced statistically significant correlations. Finally, bivariate correlations computed between OAI attribute estimates for jobs and the mean scores of job incumbent samples on corresponding tests and inventory scales revealed statistically significant correlations for 38 of 55 analyses performed. This work was performed on a sample of 1,414 jobs for which OAI factor and attribute scores and GATB test scores were available. At this time there are no data available describing the relation between GWI components and tests of ability.

While extensive efforts were expended in the development of these systems to assure reliable and accurate descriptions of jobs, there is little information relating to predictive or concurrent validity (McCormick, 1979). One of the reasons is highlighted by Trattner (1979), who writes that tests are usually validated against "whatever criteria are conveniently obtainable" (p. 118). Most often these criteria consist of subjective judgments which yield lower validities than those obtained with performance-based quantitative measures (Hunter & Hunter, 1984).

Trattner used a criterion-referenced validity approach in developing the PACE. He employed CODAP as the basis of the job analysis of three occupations and for the development of four criterion measures of job performance. These measures were work sample tests, job information tests, supervisory ratings, and supervisory rankings of workers. His 11 validity coefficients ranged from .03 to .68 with the median being .36. He concluded that "CODAP can be used very effectively in criterion related validity research ... to weight overall performance, to select research participants, and to provide excellent documentation for the relevance of criterion measures should the selection test be challenged in the courts" (p. 119). Note, however, that Trattner did not identify ability requirements, but rather used CODAP data to develop job-related tests.

The evidence of criterion-referenced validity of the ARS physical ability scales far outweighs the evidence of such validity for the other ability scales. Fleishman and Mumford (1988) report validities as high as .88 in a series of studies in which physical abilities determined from the Ability Requirements Scales were related to physical requirements of jobs, although most of "... the correlation coefficients ... lay near the mid-50s" (p. 929).

Three studies using the ARS to identify cognitive ability requirements of jobs are relevant to the question of predictive validity. The first study is reported by McAnulty (1988) in which the ARS was used to identify ability requirements for entry into the Army Initial Entry Rotary Wing (IERW) training program. Tasks most indicative of performance in the primary and instrument phases of the IERW training were identified. SME rated the type and importance of the abilities required to perform each of these tasks. Tests were then developed to measure incoming trainees on the following ARS abilities: Information Ordering, Spatial Orientation, Flexibility of Closure, Inductive Reasoning, Perceptual Speed, and Deductive Reasoning. Simple correlations between the predictor tests and overall average grade of personnel completing IERW training ranged from .24 to .44.

Fleishman (1988) summarizes results of the development of a selection procedure across 38 electric power companies. The ARS were used to identify ability requirements for 79 critical tasks performed by personnel in energy control centers. He reports "... multiple correlations as high as .41 against criteria of job performance" (p. 684).

The third study was directed at differentiating ability requirements among jobs. Jones and McAnulty (1984) employed the ARS to identify the ability requirements of tasks performed by different kinds of Army helicopter pilots arising from different mission requirements: cargo, utility, aeroscout, and attack. An earlier study cited by Jones and McAnulty (Myers, Jennings, & Fleishman, 1982), based on application of the ARS to broadly-stated tasks for these missions described in the Aircrew Training Manual (ATM), indicated that the different missions required different abilities. Using more specific tasks as the base for applying the ARS, Jones and McAnulty found small but insignificant differences of ability requirements among the missions. They concluded that the differences in results between the two studies could be due in a large part to the fact that in their study, SME rated specific tasks rather "... than broad, ambiguous ATM 'tasks'" (p. 364).

Utility

According to the literature, the number of times a taxonomy (or a job analysis system) has been employed is in a sense a measure of its usefulness or utility. Of those methods addressed here in which human abilities are derived (as opposed to learning processes, for example), the clear winners are the Fine's FJA (Fine, 1988), the JEM (Primoff & Eyde, 1988), the PAQ, and the Ability Rating Scales. McCormick (1979), McCormick and Jeanneret (1988) cites extensive use of the PAQ since 1969. Similarly, the Ability Requirements Scales have been used numerous times (Fleishman & Mumford, 1988; Fleishman & Quaintance, 1984; McAnulty & Jones, 1984). The OAI has been applied a much smaller number of times (Cunningham, personal communication, November 11, 1988), and there is evidence of only two applications of the GWI (Cunningham, Wimpee, & Ballentine, 1987; Mayfield & Lance, 1988). Lopez (1988) reports that he has used the TTAS about 100 times.

Another way to evaluate utility is the consideration of cost versus effectiveness. The effectiveness of a particular method is dependent on its reliability and validity for the purpose of use. A problem in evaluating the utility of a particular taxonomy for the purpose of identifying ability requirements is a lack of measures with regard to benefits. When predictive validity correlations are known for several methods, utility can be evaluated using a formula such as that described by Hunter and Hunter (1984). However, in the case of job analytic methods, the validity of interest is that of content or construct validity. Predictive validity is more a function of the test which was used to measure the ability. Naturally, the method which identifies the relevant abilities will affect predictive validity, in that a test based on an irrelevant ability is not expected to predict success on the job. On the other hand, if a job analysis method identifies the correct ability, but the test used to measure the ability was flawed, the predictive validity would not be a function of the job analysis method.

Because predictive validity data is lacking and not necessarily a function of the method used to identify requirements, it is very difficult to estimate benefits that would be expected with a particular method. The factors to be considered, then, would be an examination of indicators of content and construct validity, such as interrater reliability and discriminant validity of ability constructs, along with relative cost and ease of use. The cost of each method is dependent on factors such as the number of raters, type of raters, method of administration, and whether materials must be developed.

The overall effectiveness of each method is discussed in detail in section III. Effectiveness is based on factors identified through this literature review which describe overall effectiveness for the purpose of selection and factors which describe the suitability of the method for Air Force use.

Summary and Conclusions from the Literature Review

In addition to the OSM, 36 taxonomic or job analysis systems were reviewed. From this review, it is apparent that those methods categorized as Behavioral Requirements and Task Characteristic approaches do not lend themselves easily to the definition of ability requirements of Air Force specialties. While it can be argued with merit that methods from these approaches could serve as a base for determining abilities, the fact remains that the level of expertise needed to apply the methods is not readily accessible. More importantly, there is not yet a fully-developed model for linking the intervening processes or task characteristics with construct valid measures of these processes or characteristics. For this reason, methods classified under these two approaches did not warrant further attention.

Similarly, special purpose taxonomies for management positions do not appear to warrant further consideration. While these methods describe

management jobs, the kinds of information provided closely replicates that provided by more general purpose methodologies.

From the taxonomic approaches remaining, Behavior Description and Ability Requirements, seven methods remain for consideration. Two are Ability Requirements methods, Fleishman's Ability Requirements Scales and Loper's Threshold Traits Analysis System. The remaining five are the PAQ, OAI, GWI, JEM, FJA (the Department of Labor methodology is not included because of its high degree of similarity to FJA). All seven methods are for general purpose application, most of them having been used in a variety of situations. The differences lie in the way job requirements are defined (job element description versus ability identification) and the kind of inferential process involved in linking the abilities with valid tests.

Based on the evaluative criteria proposed in the literature pertaining to internal versus external validity and utility, there is little to choose from among them. Reported reliabilities (intraclass correlation, rate-rerate) are very similar and reach at least minimally acceptable levels. Coverage provided by each of them, as evidenced by the extent to which most of them has been used, seems adequate, although clearly the best and most direct coverage is provided by the ARS. All with the exception of the GWI and the OAI have been used frequently, and by analogy, one could consider the OAI and GWI to have utility, because of the extensive usage of their predecessor, the PAQ. The GWI especially is appropriate to be included, because its recent application has been to Air Force jobs (Cunningham et al., 1987; Mayfield & Lance, 1988).

These methods, however, differ in some very important ways that affect their utility for Air Force use. These issues will be elaborated and discussed in Section III.

III. EVALUATION OF TAXONOMIES FOR AIR FORCE USE

This Section contains an evaluation of the candidate taxonomies in terms of the criteria provided by the literature review and several other criteria relevant to the selection of a taxonomy of human abilities for Air Force application. There is also a discussion of the requirements for generalizing ability requirements across specialties and options for obtaining descriptive data for specialties (primarily officer) for which OSM data are unavailable. The final subsections contain a summary of the evaluation factors applied to the candidate taxonomies, conclusions, and recommendations.

Overall, four categories of candidate taxonomies (Table 5) were evaluated in terms several criteria, including the criteria derived from the literature review (Section II): Behavioral Description, Abilities Requirements, Behavioral Requirements, and Task Characteristics. The final list of candidate taxonomies, for reasons cited in Section II, came from the Behavioral Description and Abilities Requirements approaches. Five methodologies were Behavioral Description taxonomies: Functional Job Analysis (FJA) (Fine); Job Element Method (JEM) (Primoff); Position Analysis Questionnaire (PAQ) (McCormick); Occupational Analysis Inventory (OAI) (Cunningham); and General Work Inventory (GWI) (Ballentine). Two Abilities Requirements taxonomies were evaluated against the same factors: Abilities Requirements Scales (ARS) (Fleishman); and Threshold Traits Analysis System (TTAS).

Table 5. Classes of Candidate Taxonomies

<u>Classes</u>	
Behavioral Description	Occupational Survey Methodology (OSM); Functional Job Analysis (FJA) (Fine) Job Element Method (JEM) (Primoff); Position Analysis Questionnaire (PAQ) (McCormick); Occupational Analysis Inventory (OAI) (Cunningham); General Work Inventory (GWI) (Ballentine).
Abilities Requirements	Abilities Requirements Scales (ARS) (Fleishman); Threshold Traits Analysis System (TTAS) (Lopez).
Behavioral Requirements	(Gagne, Miller) Cognitive, psychomotor, memory behaviors required for achieving criterion levels of performance. More appropriate for task analyses and instructional system development than for selection.
Task Characteristics	Task versus human performance characteristics. Linkage of task characteristics to human performance requirements for selection/classification or for task performance is not designed.

In addition to the criteria for evaluating taxonomic approaches provided by the literature, four other criteria are important for selecting a taxonomy for Air Force use. These criteria include compatibility with the OSM CODAP task data base; methods for linking ability requirements and job descriptive information with construct-valid selection instruments; ease of use in the Air Force environment; and coverage of the various methodologies.

CODAP Compatibility

Because of the extensive CODAP task data base and its use in personnel and training decision making, compatibility of the taxonomic approach used by the Air Force for selection is highly desirable. Some of the seven methodologies reviewed in Section II, however, would not efficiently make use of this data base. The PAQ, OAI, and GWI clearly, since they describe jobs, would replace OSM CODAP task-job descriptions; would inhibit use of CODAP data (such as percent performing, task difficulty); and would preclude the identification of the variance of jobs that exist in a given specialty. Where OSM data are not available for AFSs, however, any of these methods could substitute. FJA requires tasks developed within the context of FJA guidelines upon which job analysts make their judgments about job requirements. JEM utilizes job elements that bear little resemblance to OSM tasks. The rating scales employed in FJA could be applied to OSM tasks. The TTAS is applied at the job level, thus OSM tasks would be of little use.

The Ability Requirements Scales would make efficient use of OSM tasks and data and, in fact, require that kind of data for application. The ARS is designed to be applied at the task level, although the scales can be applied at the job level as well. Further, a data base from which to make determinations about which tasks are to be analyzed is required for ARS application.

In summary, only the use of the ARS will take advantage of OSM tasks and associated data. Use of other methods would entail a data collection effort to obtain job descriptive information in addition to the linking of these data to ability requirements and tests of these abilities.

Linkage with Construct-Valid Selection Instruments

The linkage of job descriptive information (tasks or job elements) with construct-valid selection instruments is a crucial issue and is, at best, an inferential process. There appears to be a consensus in the literature (see, for example, Levine et al., 1988, and Wernimont, 1988) that work-oriented techniques (like the OSM) require a greater inferential leap to relate tasks to ability requirements than is required for worker-oriented techniques (like the JEM, PAQ, or GWI). This consensus, however, appears to be based on the premise that specific task data are linked directly to ability tests as opposed to an intervening process (like the ARS) which defines abilities for linkage with

tests. In this regard, Levine et al. (1988), state that problems arise when there is a failure to provide support for the inferences about ability requirements when work-oriented methods are employed. The inferential leap from job analysis to job requirements must be clear and understandable.

Whatever methodology is selected, the need for a strategy to link ability requirements of Air Force jobs to construct-valid tests of those abilities is implicit. There is evidence that the linkage can be inferential. In this regard, Wernimont (1988) writes: "Current usage, especially in Equal Employment Opportunity (EEO) and judicial commentary, seems to emphasize the inferences to be made in getting to accurate measures of some attribute needed for a given job" (p. 195). Further, Wernimont states that the linkage of human abilities with tasks or duties of a job is not necessarily done through validation strategies alone. The current trend, according to Wernimont's review, is use of a consensus of supervisors, job incumbents, or job analysts to determine the needed skills and attributes for a job. Some inference is implied at the job analysis stage, "... but one does not need to make further inferences or rely on outside expert judgment in order to determine what skills and attributes might be needed for a given job" (p. 195). In the use of this inferential approach, there seems to be a blurring between it and a content-validity strategy.

The differences among the seven methods for linking job requirements to construct-valid selection instruments are significant. Two different kinds of linking processes have been used with the PAQ and OAI. First, factor scores on the instrument for jobs are related to the mean test scores of incumbents holding the jobs, a much less direct methodology than the second process. In the second process, industrial psychologists rated the elements of each instrument according to a set of human abilities (101 for the OAI and 76 for the PAQ). Job elements for any given job are then compared to these ratings to determine ability requirements. It should be pointed out that both of these approaches have merit if either synthetic (Primoff & Eyde, 1988) or job component validity (McCormick, 1979) is an objective.

In Lopez's (1988) traits or abilities method, the FJA, and the JEM, trained job analysts are the source of the linkage. They rate jobs or tasks according to a set of human abilities defined within each of the three systems. While similar to the methodology employed by Fleishman, the traits or abilities are more generally defined.

The most direct linkage, although inferential, of job ability requirements and tests of these abilities is achieved through use of the Ability Requirements Scales. Emphasis during development of the ARS was on precise definition of human abilities for which construct valid tests existed. SMEs rate job tasks on the scales (Fleishman & Mumford, 1988) to define the job requirements.

Use of the ARS to rate ability requirements of OSM CODAP tasks resolves the problem of linking a specific, task descriptive system with selection tests. The SME ratings of the tasks on the ARS provide satisfactory inferential linkage. In addition, there is a precedence for this combination: the task-based CODAP methodology has been used in conjunction with the ARS by Levine et al. (1988).

In summary, the determination of ARS ratings for OSM tasks is a viable methodology for the Air Force. The inference of ability requirements of tasks is provided by SME judgment.

Ease of Use

The ease of use of a particular methodology relates to the number and kind of SMEs required, the amount of training needed, and the data collection effort. There are relatively small differences among the methods in their ease of use. The FJA requires trained job analysts and from 4-6 SME for 1-2 days. The analyst needs 1-3 days preparation time before meeting the SME. Trained analysts and six SMEs are required by the JEM to develop the job elements. This development requires a minimum of two sessions of 5 hours each. In use of the PAQ, analysts should receive 3 days of training. Administration requires about 4 hours. Cunningham (1988) recommends training job analysts who have college-level reading skill for 3 days. The GWI, to date, has been administered to job incumbents by mail. Administration time is about 2 hours. Fleishman and Mumford (1988) recommend use of 20 SME who receive about 3 hours of training. The TTAS requires a reviewer and five trained first-line supervisors as analysts who directly observe workers. Tasks and data which are the basis of the ARS ratings are already available in the OSM data base; therefore, no initial data collection effort to describe AFS would be required.

Use of any of the methods to describe specialties for which OSM data are unavailable could be appropriate. The GWI, however, would be at least as appropriate as any of the other generic methods for describing specialties as well as being readily available to the Air Force for administration. It is relatively easy to administer by mail. The only limiting factor for the GWI is the fact that it has not been employed in surveying officer populations, but its descriptive properties (described above) appear adequate for this purpose.

In summary, while there are small differences of ease of use among the methodologies, application of the ARS to OSM tasks would eliminate at least one data collection effort. This combination of methods seems most cost efficient for those specialties for which OSM data are available. For specialties (primarily officer) for which OSM data are not available, the GWI appears to be the most cost effective methodology.

Adequacy of Coverage of Ability Requirements

Each of the systems, in its own way, provides fairly wide coverage of its descriptive elements, and the ARS, which addresses abilities as opposed to job descriptive information, was recognized as having the most complete coverage of ability requirements (see, for example, McCormick, 1979). Nonetheless, at the risk of falling prey to the "not invented here syndrome," one is left with the feeling that none of the off-the-shelf methods is sufficient to describe all facets of Air Force jobs. First, consider that the method employed for data collection binds the user to the coverage provided by that method. In view of the resources involved in collecting and analyzing ability requirements information, it is important to assure that all of the abilities of interest are included in the taxonomy.

Second, the research supporting the cognitive abilities derived from either the behavioral descriptive or the ability requirements methods are based on tests originating from the factor analytic literature (e. g., Guilford, 1956). They were developed prior to the emergence of present computer testing capabilities and present cognitive theory research. Thus, the cognitive abilities that are defined in the taxonomies reviewed are limited to those measures which existed and which are almost entirely paper and pencil tests. Job dimensions from the PAQ, for example, have been related to the GATB, which consists of paper and pencil ability measures. The same is true for the ARS, although different tests may have been employed.

It is true, of course, that some of these paper and pencil tests have been adapted to computer administration (Fleishman, 1988), but their psychometric properties are unchanged. Is it possible, however, that the computer technology provides the potential for the definition and measurement of other cognitive or psychomotor abilities undiscovered in the earlier factor analytic research? If so, have any such abilities and tests been discovered and defined, but as yet have not found their way into the more popular taxonomies? As a simple example, it is certainly not inconceivable that such Air Force specialties as Air Traffic Controller and Weapons Controllers are characterized by requirements to visualize moving spatial objects in various configurations and at differing speeds. These requirements are not easily measured by paper and pencil tests, but one visit to a video games amusement center will clearly demonstrate the possibilities for their measurement through use of computer technology.

In this regard, Fleishman (1988) writes that computers have a unique capability for measuring certain abilities as well as having the potential for measuring human functions that cannot now be measured by printed tests. He reports successful use of tests of cognitive abilities (based on the ARS) adapted to the Apple II computer that provided such scores as variance of response time, total errors, and mean time for each move. The psychometric parameters, however, were not improved over their printed counterparts. While Fleishman forecasts the potential for using computer technology to measure such

complex cognitive processes as problem solving and decision making, he cautions that the development of such measurement must be preceded by careful, systematic research. The crucial question is whether such functions have been defined and measures developed which may be used to study enlisted and officer specialties.

The "abilities" and tests discussed in the preceding paragraphs should not be confused with the present work in cognitive science. This work, as in the Learning Abilities Measurement Program (LAMP), represents the kind of careful and systematic research called for by Fleishman. The major goal of LAMP is directed at devising new models of the nature and organization of human abilities, "... with the long-term goal of applying those models to improve current personnel selection and classification systems" (Kyllonen & Christal, 1988). There is a very fundamental difference represented by the newer approach. In Hunt's (1987) view, the conventional approaches represented by the cognitive abilities of the ARS are based on a psychometric, Euclidian representation or model of intelligence. While providing "... good summaries of the abilities tapped by paper and pencil testing" (p. 15-16), this model provides only the relative descriptions of the products of thought. The newer approach represented by cognitive science is committed to modeling the process of thinking. This approach not only shifts the focus for identifying human ability to the cognitive processes but to different measurement technology.

Hunt (1982) among others has suggested that cognitive tasks from the experimental laboratory should be exported to supplement existing aptitude tests. But as Kyllonen and Christal (1988) note, efficacy has not yet been demonstrated. At the present state of the research, the most crucial problem would appear to concern the validity strategy for linking the measures to job performance. Use of a criterion-related strategy implies individual studies for each enlisted or officer specialty, while avoiding the inferential linkage issue. If inferences are to be made to link abilities to specialty requirements, as in a job component approach, some method for the analysis of specialty or job requirements which identifies ability requirement is needed. This methodology, probably some form of job or task analysis, is not yet formalized, although Kyllonen and Christal suggest that their work in LAMP may lead to such an analytic technique in the future.

The point, however, is that the work of the cognitive psychologists has identified certain "abilities" and related measures that put new light on the study of aptitude, giving reason to explore this research to determine if there are any abilities and measures that can be adapted to analyzing enlisted and officer occupations for selection purposes. For example, Kyllonen and Christal (1988) report on an unpublished correlational study in which the newer tests were administered to a large number of subjects for whom Armed Services Vocational Aptitude Battery (ASVAB) scores were available. Results showed, as one illustration of the possibilities, that a Working Memory factor subsumes the reasoning factor from the conventional tests--individual differences in reasoning ability may be due to differences in working memory capacity. This hypothesis is based on the finding that the factor on which all of the

reasoning tests loaded highly was a Working Memory factor, the test defining it (Alpha Recoding) apparently being independent of any reasoning requirement but depending instead on working memory capacity.

On the surface, it is unclear how a factor such as Working Memory could be incorporated into a taxonomic approach like those typically employed by the more conventional methods--or whether such a factor should be. As the cognitive science matures, an entirely new and different approach to the analysis and classification of tasks will surely emerge that will capitalize on the measurement of the cognitive abilities. In the meantime, what is there to be applied from this work that is useful in the short term?

Another area of abilities involved in human task performance for which there is an apparent deficiency in taxonomies reviewed concerns social and interpersonal communication skills. The descriptions of the overall Consideration factor derived from principal components analysis of data about management jobs (Landy & Trumbo, 1988) and Page's (1988) Human Resource Development and Human Resource Management factors (from data collected by the MPDQ) clearly imply requirements for social and interpersonal communication skills in the performance of management positions. These requirements can clearly be expected to be found in officer specialties as well as some, if not all, enlisted specialties. Yet, none of the taxonomies reviewed adequately addresses this area in terms of identifying selection measures to be employed. McCormick and his followers provide job elements to describe jobs in these terms, but have not extended the work to incorporate tests of the abilities. The latest version of Fleishman's ARS (Fleishman & Mumford, 1988) does not include any social "ability" factors. If the ARS taxonomy were to be utilized, it is recommended that additional communication and/or interpersonal skills be identified and included in the taxonomy. The most effective taxonomy would be one developed for Air Force occupations. It would include a subset of ARS ability constructs, the set of perceptual and psychomotor skills which have been investigated for Air Force occupations (Siegel et al., 1980) and additional communication and interpersonal skills.

Generalizing Ability Requirements Across Specialties

Establishment of ability requirements for a large number of occupations or jobs is more efficient when a job component validity approach is employed. This approach eliminates the requirement to assess the ability requirements of each specialty independently through a rating process--that is, requirements for job components obtained through a rating or testing process of a representative sample of the total number of specialties may be generalized to other specialties having similar job components.

The development of a procedure to establish job component or generalized validity of predictors for jobs consists of these three steps: (a) some method of identifying the constituent components of jobs; (b) a method for determining

for an experimental sample of jobs the human attributes required for successful job performance with a given job component which are common to several jobs; and (c) some method for combining the estimates of human attributes required for individual job components into an overall estimate of the human attribute requirements for an entire job.

There is a model for job component validity for the PAQ and OAI. As applied by McCormick and associates (McCormick et al., 1977), job dimension scores for a sample of jobs and sets of test data for incumbents of such jobs were entered into a multiple regression model. The test data for the incumbents were used as the criterion of the importance to respective jobs of the attributes measured by the tests. The predictors were PAQ job dimension scores. McCormick argues that the relationship between job elements (or dimensions) with the mean scores of incumbents on construct-valid measures of abilities constitutes evidence that these abilities are related to job performance. A high mean ability score for incumbents of a job is evidence that the ability is required by the job.

Once a "body" of empirical evidence exists, job component validity for a given job can be shown if there is evidence of the specific job elements or job dimensions and the levels of their requirements for the particular job. After these job dimensions are defined for a job, ability requirements can be defined, and no further validation is required. Since the GATB is not used by the Air Force for selection, application of job component validity concepts based on the GATB measures and use of the PAQ (or OAI) may be questionable. The planned work reported by Cunningham (personal communication, November 11, 1988) in which GWI dimensions will be related to ASVAB scores may provide a basis for job component validity involving the use of the GWI. A question which should be explored is the extent to which the assumption that workers tend to gravitate to jobs which best suit their abilities is tenable in the Air Force worker population. In civilian jobs, workers have the opportunity to move at their discretion to the jobs they prefer. This freedom is restricted by Air Force assignment policy.

There is no model of job component validity for the ARS. Their most limiting factor for Air Force use is the lack of generality of results from the study of one occupation to another occupation, if the scales are employed as they typically have been. This limitation, however, is an artifact of the way the ARS have been used; i. e., SME rate individual tasks or groups of similar tasks for a single occupation at a time. Because task statements are very specific, there are very few tasks common across several specialties. The limitation can be overcome by a modification of the way the ARS are used by SME to analyze task requirements. Briefly, one modification would be the application of the ARS to the verbs used to describe the tasks in the job or occupation studied (for more information about the rationale for this modification, see Section IV). With a stratified, random sampling plan for the occupations analyzed in this way, ability requirements of verbs can be established for use in generalizing across other existing occupations, and for other uses like defining ability requirements for reorganizing classification

structures, for cross-training requirements, and for new or anticipated occupations (i. e., for new weapons systems). It is fascinating, in fact, to ponder the implications of the hierarchical clustering and principal components analysis of verbs on ability requirements for the creation of an action verb taxonomy of work.

There are precedents for the suggested approach. Functional job analysis is based on the assumption that task verbs can be related to levels of ability requirements. Bennett (1971) also used a verb-centered methodology for forecasting effects of automation on the future job market. In a sense, the GWI also represents this approach, many of the job elements consisting of gerunds, such as creating information, technical drawing, investigating, sawing, drilling, connecting, fastening, and joining.

Where there are no OSM data, job descriptive data can be collected using any of the generic methods, although, as pointed out, the GWI appears to be the most efficient instrument. Once the job elements for a specialty are collected, SMEs would rate these elements in terms of the taxonomic factors as Cunningham has suggested. Since the GWI elements are in themselves generic, their generalization to other specialties could be easily made.

In summary, a job component validity approach is more economical for Air Force determination of ability requirements across the many officer and enlisted specialties. While there are established models for the PAQ and the OAI, adaptation of the way the ARS is applied can provide the basis for job component validity applications to Air Force Specialties for a potentially wide variety of applications.

Descriptive Data for Specialties for Which OSM Data Do Not Exist

Since there are about 25 officer utilization fields and a few enlisted career ladders for which there are no OSM data available, it will be necessary to collect descriptive data for these specialties. It will not be possible, therefore, to apply the same methodology across all specialties. The different methods employed, however, should be as comparable as possible and the abilities derived must be from the same taxonomy. The options that could satisfy these requirements are as follows:

1. Use the OSM tasks for the enlisted and 20 officer specialties for which there are data as the basis for defining abilities from task-specific data, using an ARS-type taxonomy.

2. For the remaining specialties:

a. Initiate occupational surveys. This option would, perhaps, involve the longest period of time to completion; and further, there may be fields which cannot be surveyed successfully or cost effectively.

b. Have SME indicate ability requirements from carefully prepared summaries of the requirements of the specialties. This option is less satisfactory because of the lack of specific task statements.

c. Using SME from each of the fields, define the most representative tasks for each specialty. The SME then could rate these tasks on the an ARS-type taxonomy. This approach has been used in connection with the ARS (Landy, 1988).

d. Use one of the descriptive methods, such as the GWI, to define job requirements which SME would then rate on an ARS-type taxonomy. There are precedents for this approach, both McCormick and Cunningham having had judges rate each of the elements in their instruments on the human abilities required of them. Cunningham (personal communication, November 11, 1988) also has suggested the feasibility of rating the GWI in a similar manner.

Summary of Evaluation

This Section presents a review of the factors used to evaluate the candidate taxonomies and a brief summary of the evaluation of each of the candidate taxonomies. Appendix A contains Detailed Evaluative Data for each of the candidate taxonomies across each of the evaluation factors.

Evaluation Factors

Based on the literature review and the additional evaluative criteria, 11 factors were used to evaluate the candidate taxonomies. The factors are shown in Table 6.

Evaluation Summary

Table 7 shows a brief recapitulation of each candidate taxonomy in terms of the 11 evaluative factors.

Table 6. Evaluation Factors

-
1. Classes of Candidate Taxonomies
 2. Task-Based Versus Generic-Based Taxonomic Systems
 3. Determination of Human Ability Requirements
 4. Comprehensiveness of Taxonomy for Human Ability Determination
 5. Linkage of Human Ability Requirements to Tests of Those Abilities
 6. Reliability
 7. Validity
 8. Compatibility With OSM-CODAP System
 9. Utility: Taxonomic Employment Data
 10. Ease of Use: Applicability in USAF Environment
 11. Adequacy of Coverage of Social and Interpersonal Communication Skills
-

Conclusions

Detailed review of the OSM and the seven candidate human ability taxonomies leads to four general conclusions. First, of these methods, the Ability Requirements Scales applied against OSM tasks is the preferable approach for the Air Force, if an off-the-shelf untailored, one "best" method is chosen for defining sensory, perceptual, cognitive, psychomotor, and physical ability requirements of Air Force jobs. Second, use of an untailored, one "best" method may not capture the full range of abilities required to perform Air Force jobs. Third, methodology applied to enlisted and officer specialties will differ, because OSM data are not available for all officer and enlisted specialties. Fourth, the taxonomic approach applied should lead to the application of a job component validity approach to defining ability requirements across officer and enlisted specialties.

In regard to the first general conclusion:

1. Only one of the methods, the Ability Requirements Scales, is compatible with the OSM methodology and data base. Use of these Scales assumes the existence of specific tasks as the basis for ability ratings.

2. Techniques for linking data describing tasks to ability requirements differ. Only one of the methods, the Ability Requirements Scales, provides a predefined set of sensory, perceptual, cognitive, psychomotor, and physical ability scales. These are linked to task requirements by SME judgement. The PAQ, OAI, GWI, and TTAS are only descriptive methods and contain no predefined abilities like the ARS. Of these methods, only the PAQ and OAI have been linked to ability requirements, although the two approaches that have been used could be applied to the remaining methods. The first approach involves the judgment of experts about the ability requirements of each of the job elements

Table 7. Summary of Evaluation Data

Evaluation Factors	FJA	JEM	PAQ	OAI	GNI	ARS	TTAS
Class of Taxonomy	B.D.	B.D.	B.D.	B.D.	B.D.	A.R.	A.R.
Employs Specific Tasks	Yes	No	No	No	No	Yes	No
Employs Generic Tasks	No	No	Yes	Yes	Yes	No	No
Employs Observation	No	No	No	No	No	No	Yes
Provides Ability Taxonomy	Yes	No	No	No	No	Yes	Yes
Scope of Ability Taxonomy	3	--	--	--	--	52	16
Job Analysis-Ability Test Linkage	No	No	Yes	GATB	No	Yes	No
Reliability	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Content Validity	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kind of Criterion-Related Validity	Not Shown	Job Comp.	Job Comp.	Job Comp.	None	Crit.	Not Shown
CODAP Compatibility	Marginal	No	No	No	No	Yes	No
Utility	Many	Many	Many	Few	Twice	Many	Many
Ease of Use	Avg.	Avg.	Good	Avg.	Good	Good	Poor
Coverage:							
Social Skills	No	Yes	Yes	Yes	Yes	No	No
Interpersonal Communication	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: B.D. - Behavioral Description
A.R. - Abilities Requirement
Crit. - Criterion

in the instruments, while the second uses regression analysis in which ability test scores for job incumbents are related to the descriptive elements of their jobs.

3. If the methods are used intact, without modification, coverage of the human abilities among methods can differ. Only the ARS provides a set of predefined sensory, perceptual, cognitive, psychomotor, and physical abilities, considered by some experts as the best existing coverage. In contrast, the FJA, JEM, PAQ, OAI, GWI, TTAS and MPDQ, as descriptive methods, must be translated into ability requirements. Selection of the set of abilities (e. g., Dunnette's 10 cognitive abilities, GTAB) and ability measures to be applied is discretionary with the user and is, therefore, a function of the diligence and ability of the user.

4. There are no essential differences in validity, reliability, or utility among the seven methods.

5. Because OSM task data are available for most Air Force Specialties, use of the ARS would eliminate an initial data-collection effort for specialties for which the data are available. Use of the other methods would require an initial data collection effort followed by the collection of data to link the job descriptive elements to ability requirements.

In regard to the second conclusion, while an OSM-ARS approach appears to be most efficient of the most-used methods for defining human performance requirements, there is considerable evidence that application of a single "best" method is not necessarily in the best interest of the Air Force:

1. A single "best" method binds the user to the abilities covered by that system. The evidence in this review suggests the existence of abilities which are important for describing the human requirements of Air Force jobs but which are not included in any current taxonomy or job analysis methodology.

2. A need for a taxonomy of social and interpersonal communication abilities is apparent, especially for officer jobs and those enlisted specialties which involve important interpersonal contact, such as Air Passenger Specialists and Personnel Specialists.

3. Although only formative at this point in time, the results of recent research in cognitive psychology and in use of the computer as a testing medium have not found their way into any taxonomies. This research is providing insight about the nature and structure of human ability and appropriate measurement methods.

In regard to the third general conclusion:

1. If the OSM data base serves as the task-descriptive methodology, some differences in approach will be required to define ability requirements for some officer and enlisted specialties for which OSM data do not exist.

2. Descriptive data for these specialties is required, and of the options for gathering these data as the basis of determining ability requirements, the GWI appears adequate.

In regard to the fourth conclusion:

1. Application of a taxonomic approach to determine ability requirements is resource intensive; thus, the approach should be constructed so that job component validity can be applied from a random sample of specialties to other specialties. This approach will eliminate the requirement for data collection of SME judgments from each officer and enlisted specialty.

2. Use of the JEM, PAQ, OAI, and GWI permits the application of a job component validity strategy for defining ability requirements for occupations with similar job element structure. The generic nature of these methods allow them to be applied across jobs.

3. While there is no evidence that the ARS has been the basis of job component validity as it was applied in past studies, the Air Force CODAP task data base provides a basis for adapting the ARS to job component validity use. When the ARS are applied at the task level, as in the usual practice, abilities derived are not applicable across tasks, hence jobs. The ARS, however, can be applied at the verb (the action) level of a task. A verb-ability taxonomy applicable across jobs can be developed.

Recommendations

Since use of any one single, "best" taxonomy from among those most frequently applied to the study of the human performance requirements of jobs may inhibit the identification of the full range of these kinds of requirements in Air Force officer and enlisted jobs, a taxonomy adapted to the Air Force requirements should be developed. Such a taxonomy would be a modification of existing taxonomies and would be characterized by these features:

1. The taxonomy should be constructed so that full use of the OSM data base can be realized.

2. The taxonomy should be modeled after the ARS. Abilities comprising the taxonomy, if at all possible, should be those for which valid tests are available.

3. Abilities comprising the taxonomy should consist of those presently defined in the ARS or by the ETS Factor Reference Tests; abilities defined for social and interpersonal communication skills; and, as appropriate, abilities that have derived from advances made in automated testing technology and from cognitive science research. In final form, it should include elements that, used separately or in combination, would permit the identification of sensory, perceptual, cognitive, psychomotor, communication, or interpersonal ability requirements of Air Force jobs.

The taxonomy should be tested in a pilot study involving a small number of officer and enlisted specialties. This pilot study should address, at a minimum, the following issues:

1. The reliability of Air Force SMEs' judgments about ability requirements based on, first, action verbs which describe the work of a specialty and second, on GWI job element information.

2. The qualitative similarity of judgments from the two sources of descriptive information (i. e., OSM versus GWI).

3. The factor or cluster structure of the action verbs which are rated as well as the factor or cluster structure of the abilities identified.

4. The degree to which specialties are differentiated by their ability requirements.

5. The relationship of derived abilities with other indices of ability-apptitude requirements.

REFERENCES

- Ballentine, R.D., & Cunningham, J.W. (1981, November). Development of the General Work Inventory. Proceedings of the 23rd Annual Conference of the Military Testing Association, Arlington, VA.
- Bell, J.M. (1984). Special report: Professional military education - officer (AFPT 90-XXX-522). Randolph AFB, TX: USAF Occupational Measurement Center.
- Bennett, C.A. (1971). Toward an empirical, practical, comprehensive task taxonomy. Human Factors, 13, 229-235.
- Chambers, R.M. (1963). Operator performance in acceleration environments. In N. M. Burns, R. M. Chambers, & E. Hendler (Eds.), Unusual environments and human behavior. London: Glencoe.
- Christal, R. F. (1971). Stability of consolidated job descriptions based on task inventory survey information (AFHRL-TR-71-48, AD-734 739). Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.
- Cornelius, E.T. (1988). Practical findings from job analysis research. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 1). New York: John Wiley and Sons.
- Cunningham, J.W. (1988). Occupation Analysis Inventory. In S. Gael (Ed.), The job analysis handbook for business, industry, and government (Vol. 2). New York: John Wiley & Sons.
- Cunningham, J.W., Wimpee, W.E., & Ballentine, R.D. (1987). Some general dimensions of work in U. S. Air Force enlisted specialties. In M. Hakel (Chair), The dimensionality of work. Symposium conducted at the 2nd Annual Conference of the Society of Industrial and Organizational Psychology, Atlanta, GA.
- Dittmar, M.J., Driskill, W.E., & Weissmuller, J.J. (1987). Validation of learning difficulty against training criteria. Unpublished manuscript.
- Driskill, W.E., Keeth, J.B., & Gentner, F.C. (1981). Work characteristics - A task based benchmark approach. Paper presented at the 23rd Annual Conference of the Military Testing Association, Washington, DC.
- Driskill, W.E., & Mitchell, J.L. (1979). Variance within occupations: Jobs analysis versus occupational analysis. Paper presented at the 21st Annual Conference of the Military Testing Association, San Diego, CA.
- Driskill, W.E., Weissmuller, J.J., & Staley, M.R. (1987). Task identification and evaluation system: Interfacing task data bases. Unpublished manuscript.

- Droege, R.C. (1988). Department of Labor Job Analysis Methodology. In S. Gael (Ed.), The job analysis handbook for business, industry, and government (Vol. 2). New York: John Wiley & Sons.
- Dunnette, M.D. (1976). Aptitudes, abilities, and skills. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago, IL: Rand McNally College Publishing Company.
- Fine, S.A. (1988). Functional job analysis. In S. Gael (Ed.), The job analysis handbook for business, industry, and government (Vol. 2). New York: John Wiley & Sons.
- Finley, D.L., Obermayer, R.W., Bertone, C.M., & Muckler, F.A. (1970). Human performance prediction in man-machine systems: A technical review (CR-1614). Canoga Park, CA: The Bunker-Ramo Corp.
- Fleishman, E.A. (1972). Manual for supervisory behavior description questionnaire. Washington, DC: Management Research Institute.
- Fleishman, E.A. (1988). Some new frontiers in personnel selection research. Personnel Psychology, 41, 679-701.
- Fleishman, E.A., & Mumford, M.D. (1988). Ability requirements scales. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 2). New York: John Wiley & Sons.
- Fleishman, E.A., & Quaintance, M.K. (1984). Taxonomies of human performance. New York: Academic Press, Inc.
- Gagne, R.M. (1977). The conditions of learning (3rd ed.). New York: Holt, Rinehart, & Winston.
- Guilford, J.P. (1956). The structure of intellect. Psychological Bulletin, 53, 267-293.
- Hackman, J. R. (1970). Tasks and task performance in research in stress. In J.E. McGrath (Ed.), Social and psychological factors in stress. New York: Holt, Rinehart and Winston, Inc.
- Hemphill, J.K. (1959). Job descriptions for executives. Harvard Business Review, 37, 55-67.
- Hunt, E. (1982). Science, technology, and intelligence. In R.R. Ronning, J.A. Glover, J.C. Conoley, & J.C. Witt (Eds.), The influence of cognitive psychology on testing. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hunt, E. (1987). Toward new ways of assessing intelligence. Intelligence, 6, 231-240.
- Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternate predictors of job performance. Psychological Bulletin, 96, 72-98.

- Jones, D.H., & McAnulty, D.M. (1984). An examination of ability requirements for various rotary wing missions. In Proceedings of the 28th Annual Meeting of the Society, San Antonio, TX.
- Kantor, J.E., & Bordelon, V.P. (1985). The USAF pilot selection and classification research program. Aviation, and Environmental Medicine, 56, 258-261.
- Kyllonen, P.C., & Christal, R.E. (1988). Cognitive modeling of learning abilities: A status report of LAMP (AFHRL-TP-87-66, AD-A190 671), Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Landy, F.J. (1988). Selection procedure development and usage. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 1). New York: John Wiley & Sons.
- Landy, F.J., & Trumbo, D.A. (1980), Psychology of work behavior (rev. ed.). Homewood, IL: Dorsey Press.
- Levine, E.L., Thomas, J.N., & Sistrunk (1988). Selecting a job analysis approach. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 1). New York: John Wiley & Sons.
- Levine, J.M., & Teichner, W.H. (1973). Development of a taxonomy of human performance: An information-theoretical approach. JSAS Catalog of Selected Documents in Psychology, 2, 28.
- Lopez, F.M. (1988). Threshold traits analysis system. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 2). New York: John Wiley & Sons.
- Lumsdaine, A.A. (1960). Design of training aids and devices. In J. Folley (Ed.), Human factors methods for system design (AIR-290-60-FR-225). Pittsburgh, PA: American Institute for Research.
- Mayfield, D.L., & Lance, C.E. (1988). Development of a candidate task taxonomy for Air Force enlisted specialties. Unpublished manuscript.
- McAnulty, D.M. (1988). Pre-operational validation of new Army flight aptitude selection tests. Paper presented at the 30th Annual Conference of the Military Testing Association, Washington, DC.
- McAnulty, D.M., & Jones, D.H. (1984). An evaluation of aviator training ability requirements scale ratings. In Proceedings of the 28th Annual Meeting of the Society, San Antonio, TX.
- McCormick, E.J. (1979). Job analysis: Methods and applications. New York: AMACOM.

- McCormick, E.J., & Jeanneret, P.R. (1988). Position Analysis Questionnaire (PAQ). In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 2). New York: John Wiley & Sons.
- McCormick, E.J., Mecham, R.C., & Jeanneret, P.R. (1977). Technical manual for the position analysis questionnaire (PAQ). West Lafayette, IN: University Book Store.
- Miller, R.B. (1973). Development of a taxonomy of human performance: Design of a systems task vocabulary. JSAS catalog of selected documents in psychology, 3, 29-30.
- Mitchell, J.L., & Driskill, W.E. (1986). Optimizing integrated personnel system training decisions and development. In W. W. Tornow (Chair), State-of-the-Art Applications of Job Analysis: Integrated Personnel Systems. Symposium conducted at the 94th Annual Convention of the American Psychological Association, Washington, DC.
- Mitchell, J.L., Ruck, H.W., & Driskill, W.E. (1988). Task based training program development. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 1). New York: John Wiley & Sons.
- Mumford, M.D., Yarkin-Levin, K., Korotkin, A., Wallis, R., & Marshall-Mies, J. (1986). Characteristics relevant to performance as an Army leader: Knowledges, skills, abilities, and other characteristics and generic skills (ARI research Note 86-24). Washington, DC: US Army Research Institute for the Behavioral and Social Sciences.
- Murphy, K.R. (1988). Predictor-based taxonomy of Navy ratings: A preliminary study (NPRDC TN 88-35, AD-A193 695). San Diego, CA: Navy Personnel Research and Development Center.
- Myers, D.C., Jennings, M.C., & Fleishman, E.A. (1982). Validation of cognitive perceptual-motor tests for assigning pilots to different missions: Job analysis findings (Interim Report). Washington, DC: Advanced Research Resources Organization.
- Page, R.C. (1988). Management position description questionnaire. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 2). New York: John Wiley & Sons.
- Peterson, N.G., & Bownas, D.A. (1982). Skill, task structure, and performance acquisition. In M.D. Dunnette & E.A. Fleishman (Eds.), Human performance and productivity: Human capability assessment. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Posner, M.I. (1964). Information reduction in the analysis of sequential tasks. Psychological Review, 71, 491-504.
- Primoff, E.S., & Eyde, L.D. (1988). Job element analysis. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 2). New York: John Wiley & Sons.

- Rose, A.M., Czarnoiewski, M.Y., Gragg, F.H., Austin, S.H., Ford, P., Doyle, J., & Hagman, J.D. (1984). Acquisition and retention of soldiering skills (Contract No. 1MDA-81-C-AA01). Washington, DC: U. S. Army Institute.
- Siegel, A.I., Federman, P.J., & Welsand, E.H. (1980). Perceptual/psychomotor requirements basic to performance in 35 Air Force specialties (AFHRL-TR-80-26, AD-A093 981). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Specialized Undergraduate Pilot Training (SUPT). (AFPT 90-1XX-855) (1988). Randolph AFB, TX: USAF Occupational Measurement Center.
- Theologus, G.C. (1970). Development of a taxonomy of human performance: A feasibility study of ability dimensions for classifying human tasks (AFOSR-70-0930TR, AD-705 672). Pittsburg, PA: American Institutes for Research in the Behavioral Sciences.
- Tornow, W.W., & Pinto, P.R. (1976). The development of a managerial taxonomy: A system for describing, classifying, and evaluating executive positions. Journal of Applied Psychology, 61, 1-11.
- Trattner, M.H. (1979). Task analysis in the design of three concurrent validity studies of the Professional and Administrative Career Examination. Personnel Psychology, 32, 110-119.
- Uniform guidelines on employee selection procedures (1978). Federal Register, 1978, 43, 38290-38315.
- USAF Program Technical Training (PTT), Section VIII (1988, November). Randolph AFB, TX: DCS Technical Training.
- Wernimont, P.F. (1988). Recruitment, selection, and placement. In S. Gael (Ed.), The job analysis handbook for business, industry and government (Vol. 1). New York: John Wiley & Sons.
- Wheaton, G.R., & Mirabella, A. (1972, October). Effects task index variations on training effectiveness criteria (NAVTRAEQUCIPEN 71-C-0059-1). Orlando, FL: Naval Training Equipment Center.
- Work inventory of enlisted specialties, AFPT 90-XXX-487 (1982). Randolph AFB, TX: USAF Occupational Measurement Center.

APPENDIX
DETAILED EVALUATIVE DATA

TASK-BASED VERSUS GENERIC-BASED TAXONOMIC SYSTEMS

<u>CANDIDATE SYSTEM</u>	<u>REQUIREMENT FOR ANALYSIS BASELINE OF SPECIFIC TASKS</u>
FJA (Fine)	YES: According to FJA – specific rules
JEM (Primoff)	YES: For Job Elements according to JEM – specific rules
PAQ (McCormick)	NO : Instrument provides general descriptive data at job level
OAI (Cunningham)	NO : Instrument provides general descriptive data at job level
GWI (Ballentine)	NO : Instrument provides general descriptive data at job level
ARS (Fleishman)	YES: From any (task, task cluster, job level) source
TTAS (Lopez)	NO : Instrument provides general descriptive data at job level

DETERMINATION OF HUMAN ABILITY REQUIREMENTS

<u>CANDIDATE SYSTEM</u>	<u>DETERMINATION OF HUMAN ABILITY REQUIREMENTS</u>
FJA (Fine)	From scales: Reasoning (6 levels); mathematical (5 levels); and language development (6 levels). Task – level system.
JEM (Primoff)	From job element descriptions which are to include ability requirements.
PAQ (McCormick) and OAI (Cunningham)	Instrument includes broadly descriptive elements involving mental activities. Factor analysis, ratings of PAQ / OAI elements by experts against ability requirements, and correlational analysis of job profiles (dimensions) with GATB subtests are methods for determining human ability requirements.
GWI (Ballentine)	Instrument includes some descriptive elements involving mental activities. Planned correlational study of GWI profiles (dimensions) with ASVAB test results.
ARS (Fleishman)	Tasks rated on 52 precisely defined human ability requirement behaviorally anchored scales which are directly linked to ability tests.
TTAS (Lopez)	Job ratings (not task ratings) on 16 generally defined abilities.

COMPREHENSIVENESS OF TAXONOMY FOR HUMAN ABILITY DETERMINATION

CANDIDATE SYSTEM

COMPREHENSIVENESS OF TAXONOMY

FJA (Fine)

SME developed; descriptive task – based coverage; human ability requirements limited to 3 scales.

JEM (Primoff)

SME developed; descriptive job element coverage; human ability requirements developed for each job element.

PAQ (McCormick)

Good descriptive coverage at general level; ability requirements developed for each job element.

OAI (Cunningham)

Extensive descriptive coverage at general level; ability requirements must be derived from descriptive data which includes 3 classes of mental activities.

GWI (Ballentine)

Good descriptive coverage at general level; ability requirements must be derived from descriptive data which includes 15 "thinking" elements.

ARS (Fleishman)

Cited as best existing coverage. Task – based; SME developed; 52 pre – existing behaviorally – anchored human attribute scales linked to ability tests.

TTAS (Lopez)

Limited to 16 abilities. Job ratings (vice task ratings) on 16 broadly defined abilities, 9 of which are cognitive elements.

LINKAGE OF HUMAN ABILITY REQUIREMENTS TO TESTS OF THOSE ABILITIES

CANDIDATE SYSTEM

LINKAGE TO HUMAN ABILITY TESTS

FJA (Fine)	Test selection based on functional job analysis. (Indirect)
JEM (Primoff)	Test selection based on job element analysis. (Indirect)
PAQ (McCormick)	Test linkage to GATB and 19 commercial tests.
OAI (Cunningham)	Test linkage to GATB.
GWI (Ballentine)	Test linkage not yet shown. ASVAB linkage test planned.
ARS (Fleishman)	Laboratory development of human ability taxonomy. Construct valid tests identified.
TTAS (Lopez)	Tests selected based on TTAS analysis. (Indirect)

RELIABILITY AND VALIDITY OF TAXONOMIC SYSTEMS

<u>CANDIDATE SYSTEM</u>	<u>RELIABILITY</u>	<u>VALIDITY</u>
FJA (Fine)	.80 ±	Content (SMEs used)
JEM (Primoff)	.80 ±	Content (SMEs used)
PAQ (McCormick)	.80 ±	Content (SMEs used) Construct (GATB) Job Component
OAI (Cunningham)	.80 ±	Content (SMEs used) Construct (GATB) Job Component
GWJ (Ballentine)	.60 ±	Content (SMEs used) Job Component (ASVAB pending)
ARS (Fleishman)	.80 ±	Content (Laboratory based) Construct (Tests : abilities) Predictive (with some physical job requirements)
TTAS (Lopez)	.80 ±	Content (SME : Job level)

**COMPATIBILITY WITH OSM – CODAP SYSTEM AND UTILITY OF
TAXONOMIC SYSTEMS**

<u>CANDIDATE SYSTEM</u>	<u>OSM – CODAP COMPATIBILITY</u>	<u>FREQUENCY OF USE</u>
FJA (Fine)	Could be adopted (task based)	Many times
JEM (Primoff)	No (job element based)	Many times
PAQ (McCormick)	No (general job level data)	Many times
OAI (Cunningham)	No (general job level data)	Few times
GWI (Ballentine)	No (general job level data)	Twice
ARS (Fleishman)	Yes (task based)	Many times
TTAS (Lopez)	Could be adopted at task level from job level data	100 + times

EASE OF USE : APPLICABILITY OF TAXONOMY IN USAF ENVIRONMENT

<u>CANDIDATE SYSTEM</u>	<u>USAGE FACTORS</u>
FJA (Fine)	FJA – trained analysts. 4 – 6 SME. 1 – 2 days to develop task list. Task list must then be rated. 1 – 3 days preparation time for analyst.
JEM (Primoff)	JEM – trained analysts. 6 SME. 2 sessions at 5 hours each. Develop set of elements for each job.
PAQ (McCormick)	2 – 3 days training time for job analyst, supervisor, and incumbents. 4 hours administration time.
OAI (Cunningham)	College – level reading skill required. 3 days of training time and practice.
GWI (Ballentine)	Only job incumbents required. Mail administration. 2 hours administration time.
ARS (Fleishman)	20 SME. 3 hours training time.
TTAS (Lopez)	Requires reviewer and 5 first – line supervisors as analysts. Training required. Direct observation of job required.