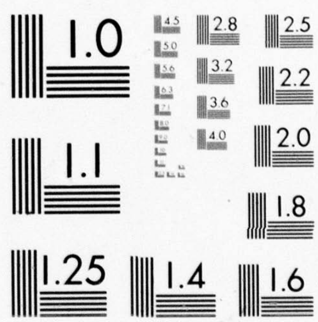
AD-A083 788 AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TX F/G 5/9
A METHODOLOGY TO EVALUATE THE APTITUDE REQUIREMENTS OF AIR FORC--ETC(U)
NOV 78 L BURTCH

UNCLASSIFIED NL

1 OF 1

AD
A083788

END
DATE
FILMED
6-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

A METHODOLOGY TO EVALUATE THE APTITUDE REQUIREMENTS OF AIR FORCE JOBS

LLOYD BURTCH
OCCUPATION AND MANPOWER RESEARCH DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
BROOKS AFB TEXAS

Presented at

20th Annual Conference of the Military Testing Association

30 Oct-3 Nov 78

Oklahoma City, OK

A METHODOLOGY TO EVALUATE THE APTITUDE REQUIREMENTS OF AIR FORCE JOBS

By
Lloyd Burtch
Occupation and Manpower Research Division
Air Force Human Resources Laboratory
Brooks AFB, Texas

I. INTRODUCTION

Aptitude requirements for entry into various Air Force career ladders are presently determined in part by the judgement of responsible personnel and in part by tradition or precedent. A precise correspondence between the aptitude scores of Air Force personnel and the aptitude requirements of Air Force jobs is extremely important since the Air Force recruits a fixed amount of talent every year and there is more demand for this talent than one might expect. There exists an additional requirement for contingency plans should the talent pool shrink or offer fewer highly talented individuals. If such shortages were to occur, which specialties could tolerate lower aptitude requirements? Which specialties could be shredded out into different job types some requiring high level talent and some low level talent? Cost effectiveness enters the picture also. Even assuming the current talent remains unchanged, it may be more cost effective to shred some specialties into jobs with varying aptitude requirements because of differences in the actual tasks performed.

More precise information about aptitude requirements will have many repercussions for the Air Force personnel system, including procurement and training. A decision to lower the aptitude entry level for a given specialty could have devastating effects on the attrition rate for the corresponding training course if no change is made in the course curriculum. For example, if an electronics course was designed for personnel with an Armed Services Vocational Aptitude Battery (ASVAB) score of E-80 or better, the existing training program is very likely to be too difficult for those with lower aptitudes. However, the aptitude level required to be successful in the training course may or may not be the same level required for success in learning how to perform the job. It is consequently possible for the Air Force to waste talent by assigning high aptitude personnel to specialties that do not require high aptitudes; and to frustrate Air Force personnel by assigning them to jobs that do not fully utilize their talents; while simultaneously neglecting other specialties in which talent is urgently needed.

The Air Force Human Resources Laboratory (AFHRL) has initiated the first systematic study to fully evaluate the aptitude requirements of Air Force specialties. The approach, ~~originated by Dr. Raymond E. Christal,~~ uses measures of learning difficulty at the task level to infer aptitude. The methodology was developed in an evolutionary manner from research documented by Fugill (1972, 1973). Christal (1973) as well as Maginnis, Uchima and Smith (1975) have further described this technology. The present paper will describe the development of task difficulty benchmark scales, their application, and will include a brief discussion of the results.

II. BASIC CONCEPTS

Task Difficulty

Task difficulty was operationally defined in terms of the time it takes to learn to perform a task satisfactorily. Based on Fugill's demonstration (1972) of high relationships between task difficulty and task aptitude ($r > .89$), this research has been conducted under the assumption that the aptitude level required to learn a job can be inferred from task difficulty, as defined above, of the tasks that make-up the job.

Benchmark Scales

A technique was required that would allow for the comparison of the learning difficulty of tasks both within and across Air Force specialties. A difficulty scale, using one or more tasks at each scale point as examples of that level of difficulty, would fill this need. Table 1 presents a simple example of such a scale. Task-anchored or benchmark scales were demonstrated to produce more reliable ratings of several task factors than did numerically anchored scales in a study by Peters and McCormick (1966). The feasibility of using task difficulty benchmark scales has been demonstrated by Fugill (1972, 1973).

Table 1. Example Benchmark Scale

Level 1 - Very Low Task Difficulty
Visually inspect batteries

Level 2 - Low Task Difficulty
Check fuse indication

Level 3 - Average Task Difficulty
Adjust transmissometer projector lamp voltages

Level 4 - High Task Difficulty
Trouble-shoot wind measuring sets

Level 5 - Very High Task Difficulty
Trouble-shoot aircraft flight control circuits

Aptitude Areas

There are four aptitude areas in the Air Force personnel testing system: general, administrative, mechanical and electronics. This research does not question the appropriateness of these areas; it is concerned with the relative order of aptitude area score requirements for specialties and jobs within each of those areas.

III. DEVELOPMENT OF BENCHMARK SCALES

Task difficulty benchmark scales have already been developed for the electronic, mechanical and general aptitude areas. The approach was similar for all scales, but the mechanical scale will be used as an example.

A general description of the scale development effort was presented by Hart (1977) at last years Military Testing Association Conference in San Antonio. The 15 specialties shown in Table 2 were selected for the mechanical scale development. These specialties are representative both of the complexity and the variety of tasks within the mechanical aptitude area.

Table 2. Mechanical Specialties
(N Task and ASVAB Cut Off)

Air Force Specialty	N Task	Mech ASVAB Cut Off
464X0-Explosive Ordnance Disposal Spec.	551	60
431X0-Helicopter Mech.	577	50
542X2-Electrical Power Production Spec.	592	50
546X0-Liquid Fuel Systems Mech.	1018	50
427X1-Corrosion Control Spec.	457	50
361X0-Outside Wire and Antenna Mech.	476	40
423X2-Aircrew Egress Systems Mech.	376	40
423X3-Aircraft Fuel Systems Mech.	297	40
426X2-Jet Engine Mech.	415	40
552X0-Carpenter	563	40
552X5-Plumber	407	40
566X1-Environmental Support Spec.	556	40
551X1-Construction Equip. Operator	927	40
427X3-Fabrication and Parachute Spec.	553	40
551X0-Pavements Maint. Spec.	927	40

Table 3. Estimates of Interrater Reliability

	Specialty	N (Rater)	R_{KK}
464X0	Explosive Ordnance Disposal Spec.	88	.96
431X0	Helicopter Mech.	100	.97
542X2	Electrical Power Production Spec.	58	.96
546X0	Liquid Fuel Systems Mech.	81	.96
427X1	Corrosion Control Spec.	43	.88
361X0	Outside Wire and Antenna Mech.	38	.93
423X2	Aircrew Egress Systems Mech.	53	.88
423X3	Aircraft Fuel Systems Mech.	26	.93
426X2	Jet Engine Mech.	83	.94
552X0	Carpenter	68	.93
552X5	Plumber	116	.97
566X1	Environmental Support Spec.	56	.94
551X1	Contruction Equip. Operator	83	.97
427X3	Fabrication and Parachute Spec.	73	.94
551X0	Pavements Maint. Spec.	72	.97

Relative ratings of task difficulty are routinely obtained in conjunction with job inventories and occupational surveys conducted by the USAF Occupational Measurement Center, Lackland AFB. These data, obtained from incumbent supervisors, are collected on all tasks in the job inventories and are provided to AFHRL for research purposes. Table 3 reflects the estimates of interrater reliability (Lindquist, 1953) and the number of raters for the 15 mechanical specialties. Using these data and the criteria outlined in Table 4, forty tasks were selected from each specialty to establish a set of 600 benchmark tasks.

Table 4. Task Selection Criteria

1. Eliminate supervisory tasks
2. Capture range of difficulty
3. Select on High Rater Agreement (Low SD)
4. Tasks performed by first termers
5. Prefer well known tasks
6. Prefer easily observed tasks
7. Face validity

In preparation for selecting the tasks from the benchmark set to represent the 25 points on the benchmark scale, a panel of mechanical experts, provided by an Air Force contractor, was asked to provide a rank-ordering of the 600 tasks. Each panel member, after accumulating detailed information on each task, provided an independent rank-order of the set of 600 tasks. The task requiring the least learning time was assigned number 1 and the task requiring the greatest learning time was assigned number 600. The estimate of interrater reliability was very high (R_{KK}=.97, N=8). This result demonstrates that a panel of work area experts can work within our definition of task difficulty, collect detailed information in the field at the task level, and provide highly reliable rank orderings of a large number of tasks selected from a given specific work area.

To address the matter of validity, the contractor's ranking data were correlated with the field supervisor's relative ratings referred to earlier. These correlations were computed using mean ranks and ratings on the forty tasks from each of 15 specialties separately; results are summarized in Table 5. These coefficients provide some substantiation of the validity of the data collection procedure, the definition of learning difficulty, and of the data itself.

Table 5. Correlations between Mean Ranks and Mean Ratings of Forty Tasks

Specialty	r
464X0	.87
431X0	.91
542X2	.87
546X0	.85
427X1	.81
361X0	.77
423X2	.83
423X3	.79
426X2	.74
552X0	.76
552X5	.57
566X1	.76
551X1	.82
427X3	.81
551X0	.73

Benchmark Task Selection

Two tasks were selected to represent each of the learning difficulty levels of the 25-point scale. A systematic procedure was developed to insure that the selected tasks represented the distribution of the mean ranks of the 600 tasks. In addition, the criteria summarized in Table 4 were again applied as appropriate. Face validity was even more important in this task selection process than it was in the prior process in as much as the tasks were to be used as examples that would anchor the various points on the scale. That is, the tasks on the mechanical scale must appear to be mechanical tasks to the extent possible.

A sample of the 50 selected tasks (two for each of 25 points) along with mean and standard deviation from the ranking process is at Table 6. The mean standard deviation for all 600 tasks was 62.8. Table 6 indicates the type of tasks selected as well as the relatively high rater agreement for most of them.

Table 6. Example Benchmark Tasks - Mechanical Scale

Level	Task Title	$\bar{X}$	SD
1	Police Grounds for Litter	1.50	.87
1	Police Open Storage Areas	3.50	1.73
5	Clean Life Preservers	26.38	13.77
5	Dig Ditches by Hand	27.00	14.41
10	Clean or Regap Spark Plugs	136.38	53.97
10	Caulk Areas Around Windows, Sinks or Bathtubs	140.63	105.52
13	Install or Replace Water Fountains	307.38	77.31
13	Disassemble or Clean Conventional Fuel Gate Valves	306.13	83.64
15	Perform Preoperational Inspections of Engine after Engine has been on long Standby	401.63	88.07
15	Install or Replace Formica on Counter-tops or Splashboards	404.13	74.44
20	Install Tail Roter Assemblies on Helicopter Aircraft	562.50	24.09
20	Read and Interpret Schematic or Wiring Diagrams	562.00	58.41
25	Troubleshoot Installed Engines	599.38	1.32
25	Troubleshoot Systems for Breaker Trip-Outs	595.38	5.20

IV. PROCEDURAL GUIDE

Accurate application of the benchmark scale requires detailed knowledge of both the task being rated and the reference tasks at each level of the scale. A procedural guide has been assembled for each scale describing the reference tasks. This guide is for the use of the panel of expert raters who will actually apply the scales. There are two parts: Part I introduces each panel member to the task of assessing learning difficulty and rating the tasks; Part II presents the 25-point scale and provides a one page description of each of the 50 tasks on the scale. This description includes the level of the task on the scale, the title of the task, the specialty from which it was selected, a narrative on any specific equipment associated with the task, a narrative describing the actual task performance, and an explanation of the skill and knowledge required to learn the task. Examples of these descriptions, taken directly from the Mechanical Procedural Guide (Hart and Pulliam, Note 1), are at Figures 1 and 2.

Figure 1. Level 10 Task Description

Level 10: CLEAN AND REGAP SPARK PLUGS (Electrical Power Production Specialist - AFSC 54350

Equipment: The task concerns gasoline engines of one or two cylinders, driving service equipment such as air compressors. These engines are part of the support equipment in an electrical power generating station.

Task Description: The task requires standard hand tools and an air blast powered spark plug cleaner which blows an abrasive against the plug base to clean insulator and electrodes. Work is performed in the power station. The mechanic removes plugs from the engine, using a socket wrench. He cleans the plug by inserting it into a hole on the cleaning machine, and pressing a valve to release a blast of abrasive against the plug base. After a few seconds he removes the plug, inspects it visually for clean ceramic, and (on some machines) inserts it in a second hold for a pressure test. Defective plugs are thrown away. He then checks the gap using a gap gauge (with feeler wires), and corrects any error by bending the outer electrode inward, using a slotted wrench which is often part of the gap-gauge handle. He puts a new plug gasket on the plug and torques the plug back in place.

Skill/Knoweldge Required: The task requires knowledge of standard hand tools, including a torque wrench. Since there is likely to be no T.O. for the engine concerned, the mechanic must know the general procedure for cleaning and gapping a plug, and that 25 foot pounds is the usual plug torque. Airmen who qualify for entry into this field usually have some knowledge of this task before their enlistment.

Figure 2. Level 25 Task Description

Level 25: TROUBLESHOOT INSTALLED ENGINES (Jet Engine Mechanic - AFSC 42652)

Equipment: This task is performed on jet engines installed on aircraft. Troubleshooting includes isolation of failure within the engine or confirming that a failure is not in the engine but some related subsystem.

Task Performance: Troubleshooting typically begins with a pilot write-up. Interpretation of these write-ups is often difficult. The isolation process depends upon the failure sympton observed. Oil leaks, which are the most common problems require that all oil be cleaned from the exterior of the engine, the engine and oil systems are isolated by attaching vibration sensors at different locations around the engine and then running the engine to look for abnormal vibration sources. Other problems such as fuel leaks, throttle rigging, fuel control, and electrical problems require coordination with other subsystem specialties to isolate the problem between the engine and related systems.

Skill/Knowledge Required: Learning troubleshooting is accomplished by exposure and is not formalized. It requires:

(a) A complete knowledge of engine operation and its interface with related aircraft subsystems.

(b) Ability to use and understand the readings of pressure gauges, vibration sensors, and heat gauges.

(c) That the mechanic be cockpit qualified to enable him to run up the engine.

(d) An ability to read and interpret the appropriate Technical Orders.

(e) Coordination with the efforts of other subsystem specialists to isolate problems in the interaction of the engine and related aircraft systems.

It is mandatory for each rater to fully absorb the contents of the guide prior to using the scale. Part I of the guide calls for a practice period of actual study and application prior to operational use of the scale.

V. APPLICATION OF BENCHMARK SCALES

The intention is to ultimately apply the scales to all available enlisted specialties in the Air Force. Data collection and analysis is underway. Because analysis is not complete, information to finalize the evaluation of the aptitude requirements in specific specialties is not yet available. Presented here is a brief discussion on how the method is to be applied.

Typically 60-70 tasks are selected from each specialty to be evaluated. These tasks will be selected using criteria similar to that used in selection of the benchmark set. The tasks will be individually studied in depth at both the technical school and at two or more operational work sites. A typical panel will be made up of 12 members with two teams of six visiting separate locations. After accumulating as much data as feasible on each task, the panel members will independently provide 1-25 point ratings of learning difficulty for all 60-70 tasks in each specialty. These ratings (for a sample of tasks within each specialty) can be used to estimate the learning difficulty of all tasks in a specialty using traditional statistical procedures for estimations.

VI. DATA ANALYSIS

The Comprehensive Occupational Data Analysis Program (CODAP) package developed by AFHRL is the data analytic tool being used in the analysis of these data. The CODAP system is ideally suited for this job. Programs are readily available to provide all necessary analysis for the project.

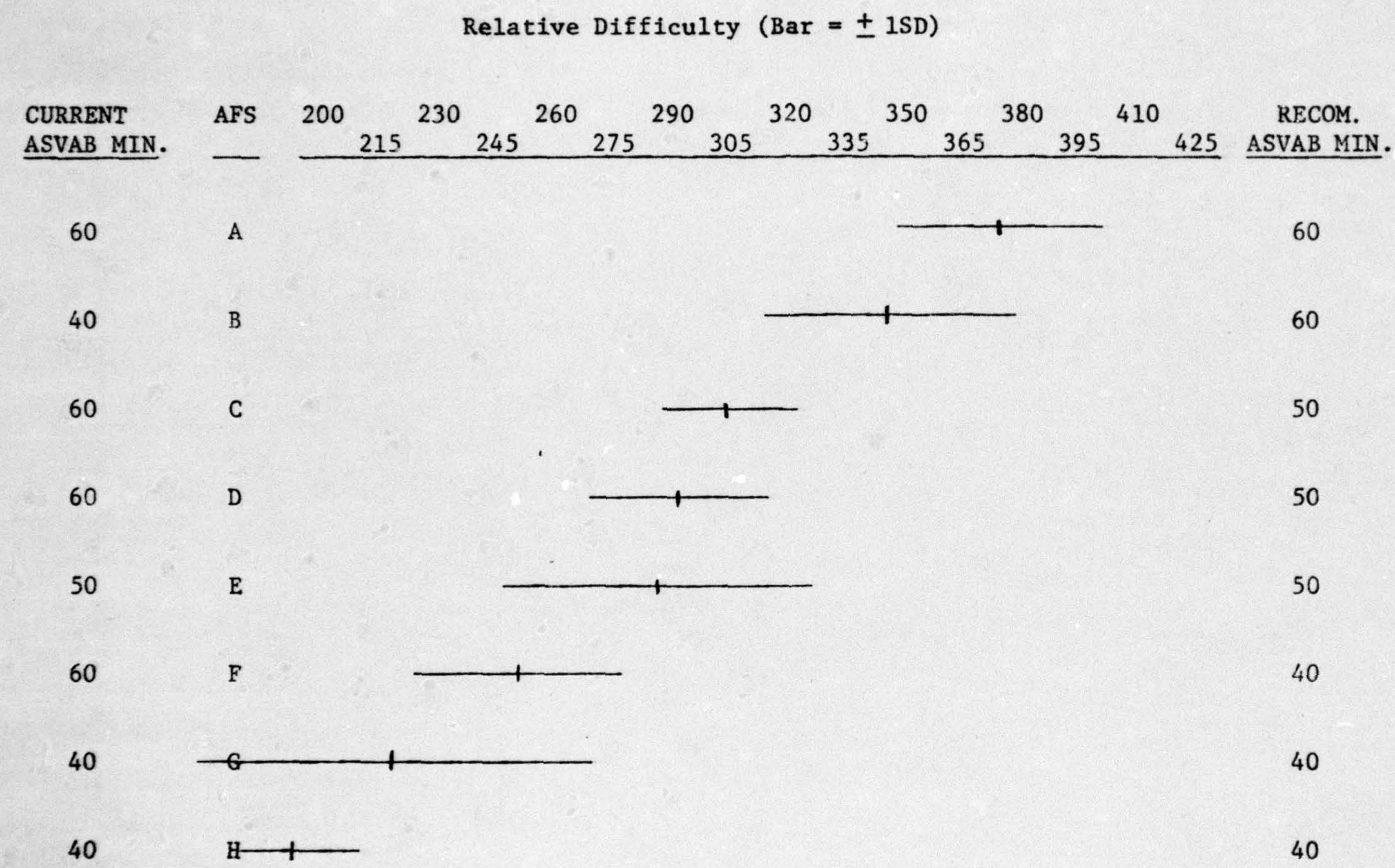
The contractor's benchmark ratings and the supervisor's relative ratings of the same 60 tasks are input to a two variable multiple regression problem for each specialty. The resulting equation is then to be applied to the supervisor's relative ratings of all tasks in the specialty. This process will result in the prediction of a 1-25 point rating mean for each task in the specialty. These predicted difficulty levels are, in turn, used as input to the CODAP system for the computation of average task difficulty for a variety of groups, and job types within each occupation. For example, the average task difficulty for first term airmen will be computed for each specialty and will be comparable across all specialties in an aptitude area. Similar computations will be made on other combinations of tasks and/or job incumbents.

VII. PRELIMINARY RESULTS

The analysis completed to date has resulted in demonstration of the efficacy of the method. Interrater agreement estimates with 12 raters rating 60 tasks from each specialty have ranged from .88 to .98. These results have convinced us that the scale, in hand with the procedural guide, can be reliably applied by knowledgeable work experts.

Some preliminary correlational analysis has been completed with positive results. Correlations between the two teams of raters have ranged from .82 to .94. Correlations between the ratings of relative difficulty and the benchmark ratings are ranging from .71 to .94. Both of these results are indicative of the validity of our methodology. Further data collection and analysis will be much more conclusive. An illustration of the planned format of the data is provided in Figure 3.

Figure 3. Relative Aptitude Requirements for First Term Jobs In 8 Specialties (Hypothetical Data)



A brief comparison of the column containing current ASVAB minimum with the column reflecting recommended ASVAB minimum indicates that there is evidence of misalignment of the aptitude requirements of these eight specialties. Specifically, Figure 3 indicates that some specialties may have a high current minimum aptitude requirement but may actually have a much lower required minimum (e.g., specialties C, D and F). The opposite is true for specialty B. Other specialties will be found to cover an extremely wide range of jobs (indicated by the length of the horizontal lines on Figure 3) suggesting that the specialty itself might be shredded out in some fashions. The information contained in Figure 3 is not based on actual data; but data of this type will soon be available on approximately 200 specialties. Changes in aptitude requirements require a total systems approach, and we do not intend to release any data in a piece-meal fashion.

VIII. CONCLUSIONS AND FUTURE PLANS

The analysis of data to date indicates that we have developed a methodology which will enable us to evaluate aptitude requirements at the task, job, and occupation level. The benchmark scale approach results in the collection of difficulty data at the task level that is comparable across all tasks within an aptitude area regardless of specialty. The results of the data analysis to date are sufficient to conclude that the total technology is based on a sound approach and analysis methodology.

There are studies in process that address the matter of longevity of the data; that is, how long will these data reflect the requirements of the specialty. Preliminary results indicate that the contractor benchmark data may be useful in assessing the learning difficulty of the specialty for several years. The difficulty scale is anchored with tasks that should not easily become obsolete because of the task selection process. First, to the extent possible, tasks were selected that were well known to mechanical workers; and second, extreme care was used in documenting each task in the procedural guides. Primarily for these reasons, it is not necessary for the tasks on the scales to even remain in active occupational task inventories to be effective. The scale will remain an effective tool as long as experts in the work area can comprehend the terminology used and the written documentation provided in the procedural guide. Not only will the scale and the benchmark data be useful in years to come, but the scales as they are will also be useful in examining the difficulty level of future tasks as they are added to job inventories. This procedure will allow the evaluation of the aptitude requirements of new specialties and/or tasks as they become a part of Air Force work.

Implementation of the results of this project is anticipated in FY 80 or 81. The primary procedure for implementation is to change the aptitude minimums as listed in AF Regulation 39-1. The results will also be implemented through the computerized job-offer system used by the AF Recruiting Service. Plans for this form of implementation are currently being prepared. We also plan to develop a total implementation package that will include complete impact analyses with recommendations for coordinated changes in the length and difficulty of Air Force resident school training courses.

There are three significant areas where cost avoidance should be achieved as a result of this research. Contingency plans for talent shortages will be available as a product of this effort. These plans will enable the Air Force to specifically plan for talent shortages in any specific specialty or across all specialties. Another product will be a more defensible position for aptitude requirements in the case of court actions. The present system, which excludes many individuals from entering Air Force jobs based on a "cut-off aptitude score," has no objective data to support its use. This research will provide data on the learning load requirements for each job. Another product will be an improved match-up of Air Force talent and job requirements. Improving this match of talent with requirements can have effects on job attitude, retention, recruiting, and training, to name just a few.

REFERENCES

Christal, R. E. The United States Air Force Occupational Research Project. AFHRL-TR-73-75. AD-774 574. Lackland AFB, TX.: Occupational Research Division, Air Force Human Resources Laboratory (AFSC), January 1974.

Fugill, J. W. K., SqdnLdr, USAF (RAAF) Task difficulty and task aptitude benchmark scales I. Mechanical and electronic career fields. AFHRL-TR-72-40. AD-754 848. Lackland AFB, TX.: Personnel Research Division, Air Force Human Resources Laboratory (AFSC), April 1972.

Fugill, J. W. K., SqdnLdr, USAF (RAAF) Task difficulty and task aptitude benchmark scales for the Administrative and General career fields. AFHRL-TR-73-13. AD-771 677. Lackland AFB, TX.: Personnel Research Division, Air Force Human Resources Laboratory (AFSC), October 1973.

Hart, F. L. Study of task difficulty using field terms and the AFHRL task anchored scales. Proceedings of the 19th Annual Conference of the Military Testing Association. San Antonio, Texas, 832-839, October, 1977.

Lindquist, E. F. Design and analysis of experiments in psychology and education. Boston: Hougton Mifflin Co., 1953, 359-361.

Maginnis, E. B., Uchima, A., and Smith, C. E. Establishing Aptitude Requirements for Air Force Jobs. AFHRL-TR-75-44 (Vol I, II & III). Vol I, AD-A023 250; Vol II, AD-A022 206; Vol III, AD-A022 250. Occupational and Manpower Research Division, Air Force Human Resources Laboratory (AFSC), October 1975.

Peters, D. L. and McCormick, E. J. Comparative reliability of numerically anchored versus job-task anchored rating scales. Journal of Applied Psychology, 1966, 50(1), 92-96.

REFERENCE NOTE

1. Hart, F. L. and Pulliam, R. Procedural guide for use of mechanical benchmark scale. Unpublished manuscript, 1977. (Available from Air Force Human Resources Laboratory, Brooks AFB, TX 78235).

ATE
LMED
-8