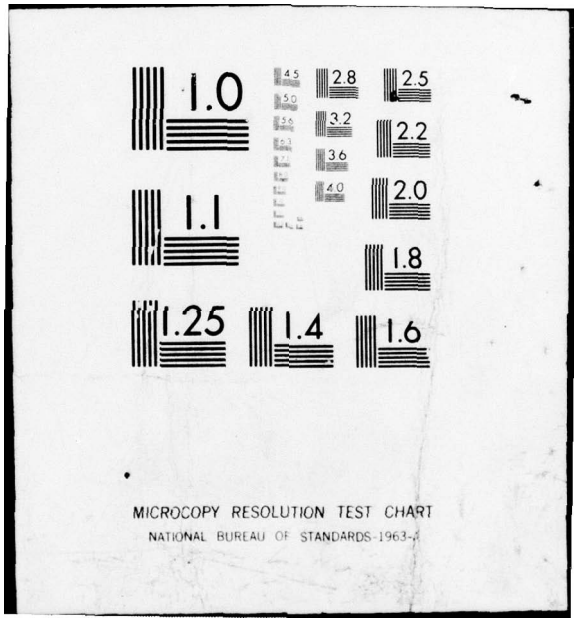
AD-A043 079 AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TEX F/G 5/9
FAMILIARITY EFFECTS ON TASK DIFFICULTY RATINGS.(U)
JUN 77 K G KOYM
UNCLASSIFIED AFHRL-TR-77-25 NL

| OF |
AD
A043079

END
DATE
FILMED
9-77
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFHRL-TR-77-25

AIR FORCE HUMAN RESOURCES LABORATORY

FAMILIARITY EFFECTS ON TASK DIFFICULTY RATINGS

By

Kenneth G. Koym

OCCUPATION AND MANPOWER RESEARCH DIVISION
Lackland Air Force Base, Texas 78236

June 1977
Interim Report for Period 1 July 1973 – 28 February 1977

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

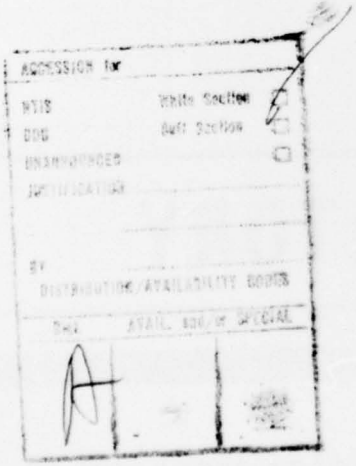
This interim report was submitted by Occupation and Manpower Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236, under project 7734, with HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

RAYMOND E. CHRISTAL, Technical Director
Occupation and Manpower Research Division

DAN D. FULGHAM, Colonel, USAF
Commander



Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFHRL-TR-77-25	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): FAMILIARITY EFFECTS ON TASK DIFFICULTY RATINGS	5. TYPE OF REPORT & PERIOD COVERED: Interim rept. 1 July 1973 – 28 February 1977
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): Kenneth G. Koym	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Occupation and Manpower Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62703F, 77340701
11. CONTROLLING OFFICE NAME AND ADDRESS: HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235	12. REPORT DATE: June 1977; 13. NUMBER OF PAGES: 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified; 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

task factors
task difficulty
task familiarity
interrater reliability
scaling context effects

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report examines interrater reliability estimates for task difficulty raters having differing levels of familiarity with rated tasks. Data were collected from 455 NCO supervisors who rated the difficulty and familiarity of 424 tasks in the Aircraft Electrical Repair career ladder task inventory. The results showed that the interrater reliability estimates (R_{kk}) decreased from .930 to .802 for six task difficulty rating scale conditions in which ratings were eliminated due to levels of familiarity. This finding suggests that little is to be gained from eliminating task difficulty ratings based upon an experienced judge's level of familiarity with tasks.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

404 415

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

PREFACE

This research was initiated under project 7734, Occupational and Career Management Research; ' task 773402, Development and Appraisal of Methods for Job Evaluation. The analyses were completed under task 773407, Development and Assessment of Methods for Determining the Requirements of Air Force Jobs; work unit 77340701, Development of Methods for Specifying Education, Training Aptitude, and Experience Requirements for Air Force Jobs.

Technical assistance and suggestions have been made by Dr. Raymond E. Christal, Squadron Leader Kenneth Goody, Dr. Joe Hazel, Mr. William J. Phalen, and Dr. Llewellyn N. Wiley. The draft manuscript was typed by Mrs. Helen Widner and proofed by Mrs. Joyce Giorgia.

TABLE OF CONTENTS

FAMILIARITY EFFECTS ON TASK DIFFICULTY RATINGS

I. INTRODUCTION

Work difficulty can be defined in several ways (Madden, 1962). Mead and Christal (1970) found task difficulty defined as "the time needed to learn to perform a task satisfactorily" to be of great benefit in predicting supervisors' job difficulty ranking policies. Equations have been developed which compare jobs within a career ladder in terms of this kind of difficulty (Koym, 1977; Mead & Christal, 1970).

Many applications have been made of task difficulty rating data. It has been possible to compare the "difficulties" of different career ladders using task difficulty benchmark scales (Fugill, 1971). Carpenter and Christal (1972) used an average task difficulty per unit time spent variable to predict the grade level of 5,485 civilian jobs. Goody and Watson (1975) also used this variable to predict task training priorities.

Improvements have been made in the reliability of rating data by eliminating ratings from widely divergent raters (having a statistically significant difference from other raters) and raters whose performance appeared to be in violation of rating instructions. An application of this technique is illustrated in recent reports by Goody (1976) and Koym (1977). However, relatively little experimental work has been reported regarding the influence of the level of familiarity on rater behavior. The current system of task evaluation makes the implicit assumption that the level of familiarity possessed does not affect the ratings provided by experienced personnel; yet, no systematic check has been made of this assumption.

This study investigates the effect of raters' familiarity on difficulty ratings assigned to tasks. It does so by asking the job incumbents to rate the degree of familiarity they possess regarding each task and the relative amount of time they estimate to be needed to learn to perform each task. The study attempts to discover if there is a systematic relationship between the levels of familiarity which raters possess with regard to a task and the level of difficulty that they assign to that task. Given a specific Air Force job inventory, what effect does familiarity with a task have on a rater's estimate of the relative difficulty of that task?

II. METHOD

The Aircraft Electrical Repair career ladder (423X0), for which a job inventory had been constructed, administered, and analyzed, was selected for the investigation. The inventory contained 424 tasks of which a subset was believed to be relatively unknown to a major part of the survey population. From the outset, task difficulty was recognized to be a complex affair (Madden, 1960, 1961, 1962; Lecznar, 1971). In this study, difficulty was understood to involve a job-oriented judgment about the properties of a task, whereas familiarity was understood to involve a personal acquaintance judgment about a task (Madden, 1960, 1962).

Seven-point relative scales were used to obtain the difficulty and familiarity ratings. Non-commissioned officers (NCO) were asked to perform two operations sequentially: (a) to rate each task for difficulty based on the time needed to learn to do the task satisfactorily using the scale ranging from "very much below average" to "very much above average," and (b) to rate each task for familiarity based on the level of acquaintance or knowledge possessed about what is involved in doing the task using the scale ranging from "none" to "complete" familiarity. For easy rater reference the above instructions and scales were printed at the top of each page of the task inventory. Appendix A contains the instructions included in the survey booklets.

Data Collection

Data collection was handled by mail via consolidated base personnel offices. Complete forms were returned by 455 job incumbents who possessed duty Air Force specialty codes (DAFSC) at the 7- and 9-skill levels.

Data Analysis Design

To test the impact of task familiarity on task difficulty interrater reliability, a series of computations was required to establish the various rating conditions of interest. These rating conditions were obtained in a three-step process by: first, identifying task difficulty ratings with seven differing degrees of rater familiarity. These seven degrees of familiarity were represented by the

familiarity scale values 7, 6 through 7, 5 through 7, 4 through 7, 3 through 7, 2 through 7, and 1 through 7 (the full scale range); second, sequentially setting tasks associated with each familiarity rating scale condition equal to 1 and those tasks associated with all other scale conditions equal to 0; third, obtaining rater-by-task cross-products between each task familiarity rating condition and the task difficulty ratings. This produced seven continuous task difficulty rating factors scaled in terms of a specific level or kind of rater familiarity.

Each of the task difficulty rating conditions was input to the CODAP REXALL interrater reliability routine (Stacey, Weissmuller, Barton, & Rogers, 1974) which applies the components of variance technique to obtain reliability estimates. The sample value for each factor was stepped up in terms of the average (k) number of raters in each rating condition using the Spearman-Brown prophecy formula. The individual difficulty ratings were adjusted (for differences in the raters' frames of reference) to a mean of 5.0 and SD of 1.0 by the CODAP REXALL routine. This was done to remove error due to systematic differences in rating tendencies and to control for familiarity context effects.

III. RESULTS AND CONCLUSIONS

Interrater Reliability Estimates

Table 1 reports reliability estimates for the task difficulty ratings obtained for various levels of rater familiarity. The intraclass correlation coefficients (R_{11}) indicate higher agreement for raters who were completely familiar (assigned values of 7) than for ratings assigned over the full-scale range (1 through 7). Raters who claimed a great deal or complete familiarity with particular tasks had higher task difficulty interrater reliability estimates than raters who claimed very little or no familiarity with the tasks.

Table 1. **Interrater Reliability Estimates for Difficulty Ratings for Different Levels of Familiarity[a]**

	Rating Scale Conditions						
	1-7[b]	2-7	3-7	4-7	5-7	6-7	7
R_{11}	.130	.142	.150	.154	.188	.208	.217
R_{kk}	.930	.928	.925	.918	.907	.884	.802
k[c]	89.5	77.5	69.6	61.1	42.2	29.0	14.6

[a]Ratings were standardized by CODAP REXALL; N Raters = 455.

[b]Full scale range.

[c]Represents one-fifth of the actual k number of raters.

However, the decrement in the Spearman-Brown (R_{kk}) values reported for the different rating scale conditions indicates that considerable loss in the stability of the mean task difficulty vectors occurs when "nonfamiliar" ratings are eliminated. The additional ratings produce higher agreement (R_{kk}) in the mean vector as is noted for rating scale condition (1 through 7) versus rating scale condition (7).

The average number of raters, k, represented in Table 1 is only 1/5 of the actual number of raters available. The lesser range was used here to bring the number of judges into a range comparable to the number of raters normally available in Air Force research studies.

As noted in Table 1, the intraclass correlation coefficients for the task difficulty ratings in this career ladder were not particularly high and probably should be explained. This may be a function of (a) the block of relatively unknown tasks which had been recently added to the task inventory, and (b) the low task performance rate present in this ladder. The low task performance rate was determined using the most recent job inventory survey previously conducted on 1581 jobs (Tartell, 1974). In light of this explanation, it may be of interest to replicate the study in other career ladders where higher reliability coefficients could be expected.

Conclusions and Implications

The results in this study were based upon ratings made by NCO supervisors who were asked to rate all 424 tasks in the job inventory survey for (a) task difficulty, and (b) task familiarity. Task difficulty rating conditions were developed for raters expressing various levels of familiarity with tasks being rated. Interrater reliability estimates were computed for these rating conditions. Although the intraclass correlation coefficients showed that the rating sets became more consistent as unfamiliar ratings were eliminated, the stability of the mean task difficulty vectors decreased when these ratings were eliminated. This increase in the internal consistency of the ratings simply does not overcome the loss in the stability of the mean vector.

While previous research on familiarity effects (Christal & Madden, 1960; Madden, 1960, 1961) has indicated that judges should have wide general experience in an area being evaluated, the results for this study have indicated that a more specific level of familiarity is not necessarily required to produce reliable jugments. It would be of interest to replicate this study in other career ladders where higher reliability coefficients exist among raters.

In keeping with the present system these results appear to support the contention that non-commissioned officer supervisors should be instructed to provide task difficulty ratings for all the tasks in job inventory surveys being administered.

REFERENCES

Carpenter, J.B., & Christal, R.E. *Predicting civilian position grades from occupational and background data.* AFHRL-TR-72-24, AD-754 966. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, March 1972.

Christal, R.E., & Madden, J.M. *Effect of degree of familiarity in job evaluation.* WADD-TN-60-263, AD-250 118. Lackland AFB, TX: Personnel Laboratory, Wright Air Development Division, November 1960.

Fugill, J.W.K. Task difficulty and task aptitude benchmark scales in exploratory study. *Proceedings of the 13th Annual Conference of the Military Testing Association,* Washington, D.C., September 1971.

Goody, K. *Comprehensive occupational data analysis programs (CODAP): use of REXALL to identify divergent raters.* AFHRL-TR-76-82, AD-A034 327. Lackland AFB, TX: Occupation and Manpower Research Division, Air Force Human Resources Laboratory, October 1976.

Goody, K., & Watson, W.J. Task factor benchmark scales for use in determining training priorities. *Proceedings of the 17th Annual Convention of the Military Testing Association,* Indianapolis, September 1975.

Koym, K.G. *Predicting job difficulty in high aptitude career ladders with standard score regression equations.* AFHRL-TR-77-26. Lackland AFB, TX: Occupation and Manpower Research Division, Air Force Human Resource Laboratory, April 1977.

Lecznar, W.B. *Three methods for estimating difficulty of job tasks.* AFHRL-TR-71-30, AD-730 594. Lackland AFB, TX: Personnel Division, Air Force Human Resources Laboratory, July 1971.

Madden, J.M. *Familiarity effects in evaluative judgments.* WADD-TN-60-261, AD-248 384. Lackland AFB, TX: Personnel Laboratory, Wright Air Development Division, November 1960.

Madden, J.M. *A further note on the familiarity effect in job evaluation.* ASD-TN-61-47, AD-263 981. Lackland AFB, TX: Personnel Laboratory, Aeronautical Systems Division, June 1961.

Madden, J.M. What makes work difficult? *Personnel Journal,* 1962, **41**(7), 341–344.

Mead, D.F., & Christal, R.E. *Development of a constant standard weight equation for evaluating job difficulty.* AFHRL-TR-70-44, AD-720 255. Lackland AFB, TX: Personnel Division, Air Force Human Resources Laboratory, November 1970.

Stacey, W.D., Weissmuller, J.J., Barton, B.B., & Rogers, C.R. *CODAP: control card specifications for the UNIVAC 1108.* AFHRL-TR-74-84, AD-A004 085. Lackland AFB, TX: Computational Sciences Division, Air Force Human Resources Laboratory, October 1974.

Tartell, J.S. Occupational Survey Report: Aircraft Electrical Repair Career Ladder, May 1974. Report available from the Defense Documentation Center and USAF Occupational Measurement Center, Occupational Survey Branch, Lackland AFB, TX.

APPENDIX A: INSTRUCTIONS

FACTOR I – TASK DIFFICULTY

DO NOT CONSIDER FACTOR II BEFORE ALL TASKS HAVE BEEN RATED ON THIS FACTOR

Instructions

STEP 1. Develop a frame of reference for rating task difficulty. Do this by scanning the entire listing of tasks. Pick out some easy tasks which can be learned in a short time, some difficult tasks can be learned only after considerable length of time, and then other tasks which fall between these two extremes. The tasks which fall at or near the middle of the range should then be used as your reference point for judging the difficulty of all tasks in the inventory. Use this reference point in completing STEP 2.

STEP 2. Estimate the time needed to learn to do each task satisfactorily compared with other tasks in the career ladder. Use the scale shown here and at the top of the difficulty column on each page to rate every task.

1. Very much below average
2. Below average
3. Slightly below average
4. About average
5. Slightly above average
6. Above average
7. Very much above average

Begin with the first task in the booklet and give each task a difficulty rating from 1 to 7; record the value opposite the task statement in the column titled "DIFFICULT." Rate every task on each page. Remember (from STEP 1) that you are comparing each task with the other tasks in the career ladder.

Record your best estimate of difficulty even on those tasks which you may believe are no longer done in your career ladder. Turn to page 1 and rate every task in this booklet for difficulty now.

FACTOR II – TASK FAMILIARITY

Instructions cont'd

STEP 3. After you have rated all tasks for task difficulty, go back to the first task in the booklet. Then, estimate the degree of familiarity which you have with each task. Consider the level of acquaintance (frequency of contact) or knowledge you possess about what is involved in doing each task. Use the scale shown here and at the top of the familiarity column on each page.

1. None
2. Very little
3. Some
4. Moderate
5. Considerable
6. Great deal
7. Complete

Begin with the first task in the booklet and give every task a familiarity rating from 1 to 7; record the value opposite the task statement in the column titled "FAMILIARITY."

Rate every task in this booklet. Estimate familiarity even on those tasks which you may believe are no longer done.

STEP 4. Write any comment you would care to make regarding the task familiarity factor or the task difficulty factor, or your ratings on either factor, on the blank pages following the last task statement. Any information which might increase the value of this study or future studies of this type will be appreciated.

STEP 5. Review the booklet to see that you have (a) completed all items of background information, and (b) given each task statement a legible rating between 1 and 7 on both the task difficulty factor and the task familiarity factor. Return this booklet to the CBPO for mailing to the AFHRL/PEOE, Lackland AFB, TX 78236.

☆U.S. GOVERNMENT PRINTING OFFICE: 1977 – 771 057/37