

# Validation and Comparison of Homogeneous and Occupational Interest Scales

**Richard R. Reilly**  
**American Telephone and Telegraph Co.**

**Gary J. Echternacht**  
**Educational Testing Service**

Occupational interest scales and homogeneous scales were developed and compared in a double cross-validation design. A sample of 3,072 Air Force personnel was randomly divided for study purposes. Homogeneous scales were developed in each half-sample using a backward item selection procedure. Occupational scales were developed by selecting items which discriminated satisfied personnel within a career field from a "men-in-general" group. A Bayesian classification procedure was then used to cross-validate the scales developed in each half-sample. The results indicated that the occupational scales were slightly superior in correctly classifying satisfied personnel versus men-in-general. However, correlations between all interest scales and job satisfaction indicated generally higher validity for relevant homogeneous scales than for occupational scales. It was concluded that where prediction of group membership or occupational choice is desired, occupational keying should probably be the preferred approach. However, when concern is with a broader range of criteria, homogeneous scales may offer more general utility.

Interest inventories, such as the Strong Vocational Interest Blank (SVIB), the Minnesota Vocational Interest Inventory (MVII), and the Kuder Occupational Interest Survey, have been widely used in guidance and vocational counseling for many years. This widespread appli-

cation has been supported by extensive research, particularly on the SVIB, which has demonstrated large and stable differences among occupations with respect to interests.

The use of these instruments within non-academic organizational settings has been limited, however. One of the reasons for this limited use may be the lack of evidence of validity for criteria other than occupational choice (e.g., Buros, 1975). Perhaps a more fundamental reason is that inventories, such as the SVIB, were intended to be of general use and, with the exception of the MVII, are oriented toward the professions.

For organizations interested in using interest assessment as a tool in career planning and placement, an obvious alternative to the published inventories is to "tailor make" an interest inventory. Such an instrument would presumably more closely reflect the characteristics of the organization and the jobs being considered. Information from job analyses could be used to develop items and scales and could be empirically developed and validated for the particular jobs of concern.

The primary purpose of the present investigation was to use such an approach for the development and validation of an interest inventory for a set of eight Air Force occupational careers. A second purpose was to compare two different methods of interest scale development.

---

*APPLIED PSYCHOLOGICAL MEASUREMENT*  
Vol. 3, No. 2 Spring 1979 pp. 177-185  
© Copyright 1978 West Publishing Co.

Methods of interest scale development have generally followed one of two principal approaches, the *occupational scale* approach and the *homogeneous scale* approach. Occupational scales are developed empirically by selecting items which discriminate significantly between members of an occupational group and a general group representative of the population of interest. Homogeneous scales, as the term implies, consist of sets of items measuring some unitary interest, such as "mechanical interest." Item analysis can also be used to eliminate items with low item-total correlations from homogeneous scales. Despite the widespread use of these two methods of scale development, few direct comparisons of occupational versus homogeneous scale development have been made, Schumacher's (1959) study being an exception. The present investigation examined the accuracy of prediction of occupational choice, as well as relationships with job satisfaction for both types of scales.

## Method

### Samples

Two all-male samples were included in the present investigation. The first sample consisted of 3,072 Air Force personnel randomly drawn from one of the following eight career fields: Weather Observer ( $N=454$ ), Radio Relay Repair ( $N=406$ ), Ground Equipment Repair ( $N=348$ ), Aircraft Maintenance ( $N=368$ ), Vehicle Repair ( $N=344$ ), Accounting Specialist ( $N=461$ ), Administrative Specialist ( $N=383$ ), and Security Specialist ( $N=308$ ). This sample was randomly divided by first grouping all personnel by career field. Then, within each of the eight fields each subject was ranked by total Airman Qualifying Examination (AQE) score. As a final step, each pair of subjects (in rank order) within each career field was considered and one of the two randomly assigned to Sample A, the other to Sample B. The result of this procedure was the formation of two half-samples in each of the eight career fields with approximately equal

AQE means and standard deviations. All personnel in the first sample had at least six, but not more than 42, months of on-the-job experience in a given career field.

A second sample consisted of 312 basic recruits who served as a "men-in-general" (MIG) group. This sample was randomly divided by a simple random assignment procedure, since AQE scores were not available.

### Inventory Development

Based on general job analysis information on Air Force career fields included in the study and a review of other occupational interest inventories, an initial pool of 400 items was developed. Each item presented an activity (e.g., "Tinker with a broken sewing machine") and required the examinee to mark a "like," "dislike," or "indifferent" response. Of the 400 items 323 were classified on an a priori basis to be in one or more of 13 major interest areas relevant to Air Force careers in general. The remaining 77 items were classified as miscellaneous.

The interest areas and number of items in each were as follows: Audiographics (18); Food Service (18); Pedagogy (12), M-Scale, a scale designed to reflect interest in activities generally related to the military (20); Leadership (13); Computational (26); Health Service (28); Scientific (34); Electronic (31); Mechanics (40); Clerical (40); Outdoors (33); and Academic (49). In the actual inventory, items were presented under each of four headings: Occupations (90 items), Work Tasks (210 items), Leisure Activities (70 items), and Desired Learning Experiences (30 items).

### Job Satisfaction Measure

Four job satisfaction scales were developed and titled "Job," "Peer," "Supervision," and "Air Force." The 16-item Job scale was designed to measure intrinsic satisfaction with the actual work activities. The 10-item Peer scale was designed to measure satisfaction with one's co-

workers. The Supervision scale contained 12 items relating to satisfaction with one's supervisor. The Air Force scale was an 8-item scale aimed at measurement of satisfaction with working conditions in the Air Force. Items were written in brief form, often consisting of only one word. Items were also stated in such a way that positive response to some items would indicate satisfaction and to other items, dissatisfaction.

The four job satisfaction scales were modifications of scales used in the Job Description Index (JDI) of Locke, Smith, and Hulin (1965). A similar approach was considered appropriate because of the relatively low reading level in the JDI and research indicating that the JDI scale scores have predictive, convergent, and discriminant validity, as well as internal consistency and stability (Robinson, Athanasiou, & Head, 1969, pp. 105-107). In addition, respondents were asked to indicate their overall degree of satisfaction on a four-choice item with the following options: very satisfied, moderately satisfied, moderately dissatisfied, and very dissatisfied.

### Analyses

*Classifying satisfied and dissatisfied personnel.* A two-step plan was initiated to assign Sample 1 individuals in each career field to satisfied within-career (SWC) and dissatisfied groups. First, within each field and in each half-sample, satisfied and dissatisfied groups were formed according to their responses to the overall job satisfaction question. Personnel responding "very satisfied" or "moderately satisfied" were classed in the satisfied group, and those "moderately dissatisfied" or "very dissatisfied" were classed in the dissatisfied group.

Second, a stepwise regression analysis was performed within each field to find the linear function of the four satisfaction scores which best separated the satisfied from the dissatisfied personnel. This linear function, or Satisfaction Index, was then used to rank all personnel

within a career field in terms of "satisfaction." The first  $N_s$  subjects with the highest scores were chosen as the satisfied group, where  $N_s$  corresponded to the number of subjects indicating overall satisfaction. This procedure resulted in the proportions of satisfied and dissatisfied personnel within each field being the same as those estimated from the overall question, but with the classifications based on the more reliable satisfaction scale scores. This procedure has some similarity to approaches described as "bootstrapping" (e.g., Dawes & Corrigan, 1974).

*Homogeneous scales.* To achieve the homogeneous scales, a backward selection procedure was used to reduce the number of items in an a priori category. All homogeneous scales were limited to a maximum of 16 items. In the case of scales with fewer than 16 items, the only items eliminated were those with negative correlations with the total scale score. The decision to use 16 items was based on the findings of Katz, Norris, and Halpern (1970), who achieved internal consistencies above .90 for interest items in 11 of 12 interest scales, using a similar item response format. The following procedure was used independently with each half-sample:

1. The correlation was computed between each item in the  $K$  items assigned to a scale and the total score based on the remaining items in the same pool.
2. The item with the lowest item-total correlation was eliminated and a new set of item-total correlations, based on the  $K-1$  remaining items, was computed.
3. The procedure was repeated until a total of 16 items remained (for the two a priori categories with less than 16 items, only those items with negative item-total correlations were eliminated).

*Occupational scales.* For the development of occupational scales the scores of personnel classified as dissatisfied were not included in the analyses. Within each career field the zero-order correlations between each interest item and a

dichotomous criterion were computed. The dichotomous criterion was scored 1 if the subject was a member of the SWC group and 0 for the comparison group, which was always Sample 2 (men-in-general). The 50 items having the highest correlations (regardless of sign) with the group membership criterion were selected from the 400 items. These 50 items served as independent variables in a stepwise regression analysis with the dichotomous criterion. Items were added until the increment in the squared multiple correlation was less than .0025. The variables, added up to this point, constituted the occupational scale. Unit weights were applied to each item for final scoring purposes. (It was necessary to reduce the search for items to comprise these scales from 400 to 50, because 400 independent variables was too large for the stepwise regression computer programs used.)

A similar procedure was used to arrive at the best linear function of a priori scale scores for separating each group of satisfied personnel from men-in-general within each career field. A stepwise regression analysis was performed, using the dichotomous criterion representing group membership. Scales adding at least .0025 to the squared multiple correlation were retained. Regression weights were rounded off to integer values for final scoring purposes.

*Cross-validation.* The occupational scales and linear functions of homogeneous scales developed in each half-sample were cross-vali-

dated on the other half-sample. For each subject in the cross-validation sample the probability of membership in either the career group or the men-in-general group was computed using the Bayesian probability function described by Cooley and Lohnes (1962, p. 138), which utilizes information regarding differences both in group frequencies and in group dispersions. For each career group the proportions of cross-validated hits and misses were compared for the homogeneous scales and the occupational scales using McNemar's test of correlated proportions (McNemar, 1962).

*Final scales.* Once the cross-validation was completed, data from the two half-samples were recombined and a final set of occupational and homogeneous scales was developed. Correlations between each of these scales and the overall Satisfaction Index (based on the combined sample for each career) were computed.

**Results**

**Classification of Satisfied Personnel**

Table 1 shows the intercorrelations and internal consistency coefficients for the four job satisfaction scales based on responses to 3,072 Air Force personnel. Also shown are the scale correlations with the overall satisfaction item. All of the correlations in Table 1 are significant. The high correlation reported between the JDI

Table 1  
Satisfaction Scale Reliabilities  
and Intercorrelations

Scale	Job	Peer	Supervision	Air Force	Overall
Job	(.840)	.348	.466	.454	.722
Peer		(.742)	.476	.351	.260
Supervision			(.884)	.379	.343
Air Force				(.597)	.373

**Note.** Internal consistency coefficients are shown in parentheses. The correlations were based on the total sample of 3,072 airmen. All correlations are significant with  $p < .001$ .

Work and People scales (Robinson et al., 1969) was not found for the Job and Peer scales in the present study, even though the two pairs of scales are similar. The estimates of internal consistency for the Job and Supervision scales were comparable to those reported for the JDI which were reported (Robinson et al., 1969) to be above .80. The internal consistencies for the Peer and Air Force scales were considerably below that standard.

The overall satisfaction question correlated much higher with the Job scale than with the other three scales, indicating that satisfaction with the intrinsic nature of the work activities which a career entails is of major importance. This result was further underscored by the consistently highest regression weight for the Job scale in the within-career regression equations used to form the Satisfaction Index.

Table 2 shows the number of personnel within each career field in each half-sample, along with the percentage classified as satisfied. The percentage satisfied within each career field was approximately the same in each half-sample. The Security Specialists had the lowest percentage satisfied by far, with 27.9% in Sample A and 20.8% in Sample B. All of the remaining career fields had at least 40% satisfied. Over all careers, the percentage satisfied was 47.9 in Sample A and 48.7 in Sample B.

### Homogeneous Scale Development

The internal consistency coefficients for the homogeneous scales ranged from .94 for Electronics to .81 for the M-Scale in Sample A. The range in Sample B was almost identical, with the lowest alpha at .82 for the M-Scale and the highest alpha at .94 for Electronics. Median alpha coefficients were .89 for Sample A and .90 for Sample B. In general, the internal consistency coefficients were slightly lower than those reported by Katz et al. (1970, p. 38) for 10 of 12 scales, but comparable to those reported by Kuder (1956, p. 21).

With the exception of the Pedagogy scale (11 items in both half-samples) and the Leadership scale (12 items in both half-samples), all scales were comprised of 16 items. Overlap between the final sets of items selected in each half-sample for each homogeneous scale was high for most of the scales. Only the Scientific (3 of 16 items overlap) and Outdoors (6 of 16 items in common) scales had less than 80% overlap for the two half-samples.

### Occupational Scale Development

The number of items selected for each occupational scale ranged from 38 to 17 in Sample A and from 27 to 18 in Sample B.

Table 2  
Number of Personnel Within Each Career Field  
and Percent Classified as Satisfied

Career	Sample A		Sample B	
	N	Percent Satisfied	N	Percent Satisfied
Weather Observer	227	47.6	227	51.5
Radio Relay Repair	203	54.7	203	56.7
Ground Equipment Repair	174	40.8	174	41.4
Aircraft & Maintenance	184	53.3	184	50.0
Vehicle Repair	172	53.5	172	54.1
Accounting Specialist	231	54.5	230	57.4
Administration Specialist	192	45.3	191	49.2
Security Specialist	154	27.9	154	20.8

Multiple correlations with the dichotomous criterion in Sample A ranged from .71 to .83. In Sample B the multiple  $R$ 's ranged from .64 to .82. Note that these multiple correlations were achieved with exact optimum weights. For cross-validation purposes, unit weights were used. Although there was some overlap in items selected in the two half-samples for the occupational scales, the amount was much less than that for the homogeneous scales.

### Cross-Validity

Table 3 shows the percentage of correct classifications achieved with the homogeneous scales and the occupational scales. In Sample A, three of the chi-square values indicated significantly more correct placements were made with the occupational scales. In Sample B, two of the chi-square values reached significance, also in favor of the occupational scales. It should be noted that the chi-square values for the different careers within a half-sample are not independent because the same men-in-general group was involved in each comparison. Although not overly strong, there was a trend for a higher percentage of correct classifications with the occupational scales.

The occupational scales also performed better relative to the "base rate" (in this case, the base rate was the  $N$  for the larger MIG group divided by the sum of the  $N$ 's for the MIG and SWC groups). Tests of correlated proportions were used to make comparisons with the base rates. In 10 of the 16 comparisons the number of hits with the occupational scales significantly bettered the base rates, as compared with 8 of 16 for the homogeneous scales. As might be expected, poorer results were obtained relative to the base rate when the SWC group was small relative to the MIG group. In the case of Security Specialist, which was the smallest SWC group, the percentage of hits achieved with the homogeneous scales was actually lower than the base rate.

### Correlations with Satisfaction

The correlations between the homogeneous scales and the Satisfaction Index are shown for each career field in Table 4. Also shown are the correlations between the relevant occupational scale and satisfaction within each career field. It should be noted that correlations for both types of scales are based on complete sample data. Since the Satisfaction Index was not involved in

Table 3  
Percent Correct Classifications Achieved with Homogeneous  
and Occupational Scales in Cross Validity Samples and  
 $\chi^2$  Values for Tests of Correlated Proportions

Career	Sample 1			Sample 2		
	Homo- geneous	Occupa- tional	$\chi^2$	Homo- geneous	Occupa- tional	$\chi^2$
Weather Observer	82.1	77.6	2.33	74.3	79.4	0.20
Radio Relay Repairman	77.8	80.5	0.36	75.6	81.9	2.69
Ground Equipment Repairman	69.0	78.8	5.13*	65.6	71.8	0.12
Aircraft & Maintenance	75.7	70.4	1.45	66.8	70.9	0.84
Vehicle Repairman	82.2	80.2	0.18	71.4	79.4	4.10*
Accounting Specialist	78.6	77.9	0.01	72.8	81.2	5.29*
Administration Specialist	68.0	78.1	7.11**	63.9	62.2	0.11
Security Specialist	56.6	69.7	7.10**	63.6	72.9	3.08

\* $p < .05$ ; \*\* $p < .01$

Table 4  
Correlations Between Interest Scales  
and Satisfaction Within Career Field

Scale	Career Field										N
	Weather Observer	Radio Relay Repair	G. Eqpt Repair	Aircraft Maint.	Vehicle Repair	Account Spec.	Admin. Spec.	Security Spec.			
Audiographic	.14**	.12**	.04	.00	.01	.06	.09*	-.03			
Food Service	.02	.06	.07	.07	.01	.05	.24**	.03			
Pedagogy	.14**	.12**	.07	.06	.03	.10	.18**	.00			
M-Scale	.21**	.19**	.12*	.14**	.07	.12**	.13**	.16**			
Leadership	.14**	.15**	.11*	.11	.10*	.15**	.23**	.05			
Computational	.21**	.15**	-.01	.02	.06	.57**	.21**	.04			
Health Service	.02	.13**	-.03	.00	.01	.02	.12*	.01			
Scientific	.15**	.16**	-.05	.01	-.02	.06	.11*	-.05			
Electronic	.17**	.27**	.16**	.08	.09*	.14**	.05	-.11			
Mechanical	.15**	.16**	.24**	.13**	.18**	.05	.11*	-.04			
Clerical	.24**	.07	.09*	.05	.10*	.35**	.36**	.10*			
Outdoors	.06	.10*	.13**	.08	.08	-.03	.02	-.09			
Academic	.00	.07	.01	-.01	.02	-.03	.20**	-.05			
Occupational	.19**	.10*	.20**	.07	.18**	.20**	.26**	.21**			
N	454	406	348	368	344	461	383	308			

\*p<.05; \*\*p<.01

the item selection for the homogeneous scales and was only indirectly involved in item selection for the occupational scales, correlations with the Satisfaction Index should not be overestimated due to capitalization on sampling error.

With the exception of Aircraft Maintenance, all of the occupational scales had significant correlations with satisfaction. In most cases, however, a logically appropriate homogeneous scale had a higher correlation with satisfaction within a career field. The Electronic scale correlated .27 with the Satisfaction Index within the Radio Relay Repair career, for example, as compared with an  $r$  of .10 for the relevant occupational scale. For Accounting Specialists, the Computational and Clerical scales correlated .57 and .35, respectively, as compared with .20 for the Accounting occupational scale. In only one career, Security Specialist, did the occupational scale have the highest correlation (.21).

Some of the homogeneous scales which were developed for general use, such as Academic and Outdoors, were nearly orthogonal to satisfaction for personnel in the eight career fields chosen. Interestingly, the M-Scale, which was designed to measure interest in militarily related activities (e.g., "March in a parade"), was significantly related to satisfaction in seven of the eight careers. In general, it might be concluded that the slight superiority of occupational scales for discriminating SWC personnel from MIG was reversed for the prediction of satisfaction.

### Discussion and Conclusions

The results of the present study do not clearly support either the homogeneous or the occupational method of scale development as better. Occupational scales were better in terms of predicting whether a subject was in the MIG or the SWC group. The homogeneous scales, on the other hand, had generally higher correlations with satisfaction. Both types of scales had some utility for both criteria.

With respect to the prediction of occupational choice, the present results are consistent with

the findings of Schumacher (1959), who found cross-validated empirical keys to be more accurate than homogeneous scales in classifying male senior medical students into post-graduate study areas. The large number of studies done with the SVIB (Campbell, 1971) further attests to the usefulness of empirical keys for categorical types of criteria. It is worth noting, however, that even with a Bayesian procedure to optimize cutting scores, the hit rate for Security Specialist was lower than the base rate. Schmidt (1974) has discussed the consequences which a relatively small occupational group has for base rates and utility of the SVIB. In the present study, all hits and misses were considered equal. Attaching specific values to different outcomes (e.g., false positives) could have produced a significantly different picture with respect to the utility of the scales.

The more general utility of occupational scales for criteria other than career choice may be limited. In the present study, correlations with satisfaction were relatively low, and lower than the correlations between satisfaction and logically appropriate homogeneous scales. As pointed out earlier, there is little evidence that occupational scales are predictive of other criteria, such as tenure. The explicit and incidental selection of SWC group members may be one factor which works to lower the usefulness of occupational keys for other criteria. Reilly and Echernacht (in press) have shown how the empirical item selection procedures used in occupational keying tend to favor items which are more highly correlated with explicit selectors than with satisfaction. Homogeneous scales, since they proceed from a logical or theoretical base, would appear to have a greater likelihood of relating to a wider range of criteria.

It is concluded that where predicting group membership or occupational choice is of chief concern, occupational keying should probably be the preferred approach. However, when concern is with a broader range of criteria, homogeneous scales would seem to be a better approach.



### References

- Buros, O. K. *Vocational tests and reviews*. Highland Park, NJ: Gryphon Press, 1975.
- Campbell, D. P. *Handbook for the Strong Vocational Interest Blank*. Stanford, CA: Stanford University Press, 1971.
- Cooley, W. M., & Lohnes, P. R. *Multivariate procedures for the behavioral sciences*. New York: Wiley, 1962.
- Dawes, R. M., & Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, *81*, 95-106.
- Katz, M. R., Norris, L., & Halpern, G. *The measurement of academic interest measures* (Research Bulletin RB 70-57). Princeton, NJ: Educational Testing Service, 1970.
- Kuder, G. F. *Kuder Preference Record Vocational Form C*. Chicago: Science Research Associates, 1956.
- Locke, E. A., Smith, P. C., & Hulin, C. L. *Cornell studies of job satisfaction: V. Scale characteristics of the Job Description Index*. Ithaca, NY: Cornell University, 1965.
- McNemar, Q. *Psychological statistics* (3rd. ed.). New York: Wiley, 1962.
- Reilly, R. R., & Echternacht, G. J. Problems with the criterion keying approach to occupational interest scale development. *Educational and Psychological Measurement* (in press).
- Robinson, J. P., Athanasiou, R., & Head, K. B. *Measures of occupational attitudes and occupational characteristics*. Ann Arbor, MI: The University of Michigan, Institute for Social Research, Survey Research Center, 1969.
- Schmidt, F. L. Probability and utility assumptions underlying use of the Strong Vocational Interest Blank. *Journal of Applied Psychology*, 1974, *59*, 456-464.
- Schumacher, C. F. A comparison of three methods for keying interest and personality inventories (Doctoral dissertation, University of Minnesota, 1959.) *Dissertation Abstracts International*, 1959, *20*, 370. (University Microfilms No. 59-2371)

### Acknowledgement

*This research was partially supported by the Personnel Research Division, Air Force Human Resources Laboratory, USAF, under Contract No. F41609-72-C-0030.*

### Author's Address

Dr. Richard R. Reilly, American Telephone and Telegraph Company, 295 N. Maple Avenue, Room 6126F3, Basking Ridge, NJ 07920.