ACT RESEARCH REPORT

No. 76

76

July 1976

ALTERNATIVES FOR VALIDATING INTEREST INVENTORIES AGAINST GROUP MEMBERSHIP CRITERIA

Dale J. Prediger

PUBLISHED BY THE RESEARCH AND DEVELOPMENT DIVISION

THE AMERICAN COLLEGE TESTING PROGRAM

P. O. BOX 168, IOWA CITY, IOWA 52240

ALTERNATIVES FOR VALIDATING INTEREST INVENTORIES AGAINST GROUP MEMBERSHIP CRITERIA

ABSTRACT

Two approaches to interest inventory validation are considered. The choice between the two depends on the use being validated. The first approach assumes that interest inventories are to be used in predicting which occupations counselees will enter or prefer. The second assumes that interest inventories are to be used in suggesting occupations for counselees to consider on the basis of compatibility of interests. Validation of these two uses of interest inventories requires different treatments of criterion group base rates. As illustrated by data drawn from a published study, the two approaches to validation can produce substantial differences in reported hit rates. Such differences may be found in any study validating group membership predictions if criterion group sizes vary greatly.

ALTERNATIVES FOR VALIDATING INTEREST INVENTORIES AGAINST GROUP MEMBERSHIP CRITERIA

Dale J. Prediger[1]

The criterion-related validity of an interest inventory is frequently reported in terms of its ability to identify accurately members of various occupations or occupational preference groups. Essentially, the validation process involves comparison of group membership predictions based on interest scores with actual group membership. (For purposes of discussion, the specific prediction procedure—e.g., high-point code, clinical intuition, maximum likelihood classification incorporating antecedent probabilities—is not directly relevant.) Once group membership predictions have been made, the number of correct predictions is tallied for each of the criterion groups. At this stage of the validation process, researchers face a crucial decision about how to treat variations in the size (base rates) of the criterion groups.

In previous discussions of the validity of group membership predictions (e.g., see Meehl & Rosen, 1955), it has usually been assumed that the goal of assessment is to maximize overall "hit rate" (the overall accuracy of the predictions). In some instances, however, the goal of assessment may be to maximize the hit rate within each of the criterion groups. The manner in which criterion group base rates are used in validity analyses reflecting these two goals can produce substantial differences in hit rates obtained for the same study. This paper examines two alternatives to the use of base rates in validating psychological assessments against group membership criteria. The application and consequences of each alternative are illustrated by examples drawn from the field of interest measurement.

Rationale for Validation Alternatives

The validity of a measuring instrument depends on the purposes for which it is used. Hence, before studying validity, one must ask "validity for what?" Interest inventories are commonly used to suggest occupational options for counselees to consider. Yet, the validity of an inventory is sometimes determined by its ability to predict future occupational preferences or occupational entry. As Berdie (1970) has noted, few counselors are interested in predicting whether a counselee will enter (or prefer) occupation A or occupation B. Hence, validity data for this use of interest inventories may not be relevant to the intended use. Furthermore, interest inventories with high validity for predicting occupational preference or entry may produce harmful side effects. Some of the reasons are discussed below.

The "Will-Prefer-or-Enter" Criterion

When predicting the occupations persons will prefer or enter, one must take into account the nature of occupational preference and employment distributions. That is, if an interest inventory is to provide accurate predictions of eventual employment, the predictions must accurately reflect the size of each occupational criterion group. To the degree that group membership predictions depart from group base rates, the inventory's predictive accuracy will be lowered.

[1]The author is grateful to Robert Brennan, Senior Research Psychologist at ACT, for suggesting use of the terms "weighted (and unweighted) average hit rate" to distinguish between the two approaches to validation.

Interest inventories which predict that males and females will enter or prefer occupations in the same proportions they have in the past should have high hit rates under this approach to validation. For a multitude of reasons (e.g., social expectations, local labor market needs, the contingencies of life), people will continue to state preferences for and enter traditional occupations. Unfortunately, the occupational preference and employment distributions of males and females are highly divergent (Gottfredson, Holland, & Gottfredson, 1975; Prediger, Roth, & Noeth, 1974). Since the predictions used in validation studies are based on the scores received by counselees, the occupational options suggested to counselees will reflect the same divergencies as the predictions, an unfortunate side effect of this approach to validation.

The "Should-Consider" Criterion

The alternative approach to the use of occupational preference and membership as criteria in validating interest inventories assumes that the purpose of interest inventories is to identify career options for counselees to consider rather than to predict the occupations counselees will prefer or enter. To achieve the former objective, an interest inventory must assess the correspondence between a counselee's interests and the interests associated with various occupational groups, regardless of the group base rates. If Cindy's interests are compatible with engineering, one would suggest that she, and others like her, consider engineering even if this lowers the accuracy with which occupational entry is predicted. The emphasis is on "should consider," not "will enter or prefer." In the course of career exploration, Cindy should find out that there are relatively few women engineers but that the situation is changing. These facts may play a role in her career decisions: they should not influence her interest score report.

Studies following this approach to interest inventory validation will treat occupational criterion groups (e.g., preference groups) as if they were of equal size. One would expect an interest inventory to suggest engineering to a large proportion of criterion group members in engineering, nursing to a large proportion of nurses, retail sales to retail sales clerks, horse shoeing to horse shoers, and so on for each of the criterion groups available. It does not matter that there are relatively few horse shoers in comparison to retail sales clerks. The question asked in the validation analysis is, "What proportion of the members of *each* criterion group would have been asked by this interest inventory to look into their occupation?" Stated differently, the question is, "What is the hit rate for each criterion group?" A high hit rate depends on an inventory's ability to differentiate the criterion groups and thus minimize the misassignment of members of each of the groups.

In this approach to validation, an interest inventory does not have to suggest retail sales to more counselees than those to whom horse shoeing is suggested because there are more retail sales clerks than horse shoers. "Predictions" are simply based on whichever criterion group a person resembles most. No premium is placed on providing interest score distributions that parallel preference or employment distributions. The proposed validation strategy recognizes that, for a number of very practical reasons, many persons may not enter the occupations suggested ("predicted") by an interest inventory.

How Choice of Criterion Affects Career Guidance

The following example may bring into sharper focus the practical implications of the two approaches to validation. Suppose that in a society built on the caste system, an interest inventory were designed to have high validity in predicting occupational entry. The inventory would suggest few if any occupations that were not traditional for a person's caste. To do otherwise would lower its validity. On the other hand, suppose an inventory were designed to identify occupational options compatible with a person's interests—regardless of the proscriptions of this society. Such an inventory might suggest many occupations not traditional for members of various castes. As a result, it would be a poor predictor of occupational entry. Yet, it might do an excellent job of determining occupational compatibility. Even in a time of social change, the score reports for such an inventory might be unsettling. But they could provide beneficial information, both to the individual and to the society.

Although useful in some types of research, interest inventories designed to predict which persons will prefer or enter a given occupation present special problems for career counseling. In effect, the rationale underlying such inventories says, "Cindy may have interests like an engineer's and Mike may have interests like a nurse's. But few females or males are likely to enter those nontraditional occupations. So let your predictions (and career guidance) take into account the relative numbers of males and females who have entered

various occupations in the past. In the long run, you'll obtain a higher hit rate and your inventory will appear to be more valid." When used in career counseling, such inventories will reinforce the society's occupational sex-role stereotypes and thus further institutionalize the channeling. Although such inventories may appear to have higher validity than inventories designed to report occupational options compatible with a person's interests, regardless of the base rates, this may be true *only* if one's purpose in assessing interests is to predict the occupations counselees will enter (or prefer). Prediger and Cole (1975) provide an extended discussion of this topic as it applies to career counseling and nontraditional occupations for males and females.

Implications for Validation Procedures

It should be apparent from the above discussion that the essential difference in the two approaches to interest inventory validation lies in the way in which criterion group base rates are treated. Two basic options are described below.

Option 1: In determining predictive accuracy, use the weighted average hit rate.

This option is frequently chosen by default. When criterion group size is "ignored" (i.e., when "hits" are simply totaled across the groups), the relative sizes of the criterion group samples determine the weighting. Hence, when Option 1 is followed, the overall hit rate is a *weighted* average hit rate. That is, the hit rates for each criterion group are weighted according to the criterion group base rates. Option 1 is preferable when validating against an occupational entry criterion. As noted later, however, other bases for differential weighting might be more appropriate when this criterion is chosen.

Option 2: In determining predictive accuracy, use the unweighted average hit rate.

This option gives each of the criterion groups equal importance in determining predictive accuracy. The objective, in effect, is to maximize the level of predictive accuracy within *each* of the criterion groups. Hence, the overall hit rate is the *unweighted* average hit rate for the criterion groups. The hit rate for a large criterion group cannot overwhelm the hit rates for several smaller groups, as in Option 1. Attention is drawn to predictive effectiveness within each of the criterion groups. Option 2 would appear to be appropriate for validating an inventory designed to suggest occupations that counselees might want to consider.

The Two Approaches to Validation Illustrated

Application of the two options to recent research results will show that the consequences of their use are far from academic. In a study comparing predictions of occupational preference obtained from Self-Directed Search (SDS) raw scores and normed scores, Gottfredson and Holland (1975) conclude that when normed scores were used with college women, "predictive validity decreased greatly" (p. 32). Criterion group hit rates in the Gottfredson-Holland study are summarized in Table 1. Holland's occupational typology is used to identify the criterion groups, and data for the two samples of college women in the study are combined to increase criterion group size. Predictions for both raw scores and standard scores (same-sex norms) are based on high-point code (i.e., highest score) for the six scales in Holland's SDS.

Results and Implications

Table 1 shows that when Option 1 is chosen, SDS raw scores produce a weighted average hit rate that is twice as large as the hit rate for standard scores. Hence, results for Option 1 support the use of raw score reports of interests, the usual procedure for the SDS (Holland, 1972). If Option 2 is chosen, however, the overall hit rates for SDS raw scores and standard scores are essentially the same. On the basis of the equivalent hit rates obtained under Option 2, one might conclude that other factors should take priority when one decides between raw-score or normed-score reporting procedures. Psychometric theory, for example, favors normed score reports. In addition, it has been shown that widely divergent career options are suggested to

TABLE 1

Variation in Hit Rates Produced by Two Validation Options

Occupational preference group	Sample size[a]	Base rate	Hit rate: Raw scores	Hit rate: Standard scores
Investigative	139	14%	45%	52%
Artistic	117	12	49	46
Social	657	67	73	24
Enterprising	48	5	4	21
Conventional	24	2	29	54
Weighted average hit rate (Option 1)			62	31
Unweighted average hit rate (Option 2)			40	39

Note. Data for 432 women in a state liberal arts college and 557 women in a state university are summarized from a study reported by Gottfredson and Holland (1975).

[a]Data reported for the "Realistic" group were not analyzed because of an inadequate N of 4. Sample sizes for the Conventional and Enterprising groups are smaller than would be desired in a well designed study.

males and females by SDS raw scores (Holland, 1972; Gottfredson, Holland, & Gottfredson, 1975). In contrast, standard scores (based on same-sex norms) suggest similar career options (Gottfredson, et al., 1975; Prediger & Hanson, 1974). In the face of equal criterion-related validity, some counselors might prefer to use raw scores while others might prefer to use normed scores. The consequences of this choice for counselees are substantial.

Discussion

In this study, the difference in results obtained for Options 1 and 2 is primarily due to the large size of the Social Group. The high raw score hit rate for this group, in combination with its size, makes a major contribution to the hit rate obtained under Option 1. However, the results would have been quite different if the Conventional Group, for example, had been the largest group—a clear illustration of the influence of group size on outcome. In this respect, Option 2 provides a more stringent index of criterion-related validity. That is, the results for a large criterion group cannot overwhelm the results for the smaller groups.

It is interesting to note in passing that one could achieve a hit rate of 67% under Option 1 simply by predicting membership in the largest group (the Social Group) for everyone. The weighted average hit rate for SDS raw scores was only 62%. This is just another instance of the base rate problem (Meehl & Rosen, 1955) resulting from application of the best a priori strategy (Cronbach & Gleser, 1965). Under Option 2, the hit rates for raw scores and standard scores "beat the base rates."

Crucial Questions

Some may not agree that Option 2 is the appropriate procedure for interest inventory validation. However, it should be clear that one must determine the purpose of assessment before deciding how to use group base rates in any study validating group membership predictions. It is not

enough to say that the purpose of psychological science is to predict behavior. One must first determine which behavior it is appropriate to predict.

To allow the relative sizes of criterion groups to influence the results of predictive validity studies (Option 1), does not "let the chips fall where they may." A true indication of the weighted average hit rate of a measure cannot be obtained unless population base rates are used instead of the base rates for the samples that happen to be at hand (Meehl & Rosen, 1955). Hence, criterion group base rates must be adjusted to reflect population base rates.

Determination of the population base rates requires answers to some difficult questions, however. For example, when validating an interest inventory via Option 1, should base rates be determined by number of workers per occupational category or by current employment needs? If the latter is chosen, should national or local needs be used? What about projected needs in 5 years? Clearly, a value judgment is involved. The decision to use Option 2 also involves a value judgment. In both instances, one must carefully examine the purpose of the measure being validated.

References

Berdie, R. F. Vocational choice studies. *Personnel and Guidance Journal,* 1970, *48*, 604-605.

Cronbach, L. J., & Gleser, G. C. *Psychological tests and personnel decisions.* Urbana: University of Illinois Press, 1965.

Gottfredson, G. D., & Holland, J. L. Vocational choices of men and women: A comparison of predictors from the Self-Directed Search. *Journal of Counseling Psychology,* 1975, *22*, 28-34.

Gottfredson, G. D., Holland, J. L., & Gottfredson, L. S. The relation of vocational aspirations and assessments to employment reality. *Journal of Vocational Behavior,* 1975, *7*, 135-148.

Holland, J. L. *Professional manual for the Self-Directed Search.* Palo Alto, Calif.: Consulting Psychologists Press, 1972.

Meehl, P. E., & Rosen, A. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin,* 1955, *52*, 194-216.

Prediger, D. J., & Cole, N. S. Sex-role socialization and employment realities: Implications for vocational interest measures. *Journal of Vocational Behavior,* 1975, *7*, 239-251.

Prediger, D. J., & Hanson, G. R. The distinction between sex restrictiveness and sex bias in interest inventories. *Measurement and Evaluation in Guidance,* 1974, *7*, 96-104.

Prediger, D. J., Roth, J. D., & Noeth, R. J. Career development of youth: A nationwide study. *Personnel and Guidance Journal,* 1974, *53*, 97-104.

ACT Research Reports

This report is Number 76 in a series published by the Research and Development Division of The American College Testing Program. The first 26 Research Reports have been deposited with the American Documentation Institute, ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. 20540. Photocopies and 35 mm. microfilms are available at cost from ADI; order by ADI Document number. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress. Research Reports No. 27 through No. 59 have been deposited with the National Auxiliary Publications Service of the American Society for Information Science (NAPS), c/o Microfiche Publications, 440 Park Avenue South, New York, New York 10016. Photocopies and 35 mm. microfilms are available at cost from NAPS. Order by NAPS Document number. Advance payment is required. Beginning with No. 60, the ACT Research Reports have been abstracted by the Educational Resources Information Center (ERIC) Clearinghouse System. Printed copies of most of the reports may be obtained from ACT Publications, P.O. Box 168, Iowa City, Iowa 52240, at a cost of $1.00 per copy. Check or money order must accompany request.

The reports in this series published since October 1970 are listed below. A complete list can be obtained from ACT Publications.

No. 37 *Practices and Outcomes of Vocational-Technical Education in Technical and Community Colleges,* by T. G. Gartland, & J. F. Carmody (NAPS No. 01441; photo, $6.80; microfilm, $2.00)

No. 38 *Bayesian Considerations in Educational Information Systems,* by M. R. Novick (NAPS No. 01442; photo, $5.00; microfilm, $2.00)

No. 39 *Interactive Effects of Achievement Orientation and Teaching Style on Academic Achievement,* by G. Domino (NAPS No. 01443; photo, $5.00; microfilm, $2.00)

No. 40 *An Analysis of the Structure of Vocational Interests,* by N. S. Cole, & G. R. Hanson (NAPS No. 01444; photo, $5.00; microfilm, $2.00)

No. 41 *How Do Community College Transfer and Occupational Students Differ?* by E. J. Brue, H. B. Engen, & E. J. Maxey (NAPS No. 01445; photo, $5.50; microfilm, $2.00)

No. 42 *Applications of Bayesian Methods to the Prediction of Educational Performance,* by M. R. Novick, P. H. Jackson, D. T. Thayer, & N. S. Cole (NAPS No. 01544; photo, $5.00; microfilm, $2.00)

No. 43 *Toward More Equitable Distribution of College Student Aid Funds: Problems in Assessing Student Financial Need,* by M. D. Orwig (NAPS No. 01543; photo, $5.00; microfilm, $2.00)

No. 44 *Converting Test Data to Counseling Information,* by D. J. Prediger (NAPS No. 01776; photo, $5.00; microfiche, $2.00)

No. 45 *The Accuracy of Self-Report Information Collected on the ACT Test Battery: High School Grades and Items of Nonacademic Achievement,* by E. J. Maxey, & V. J. Ormsby (NAPS No. 01777; photo, $5.00; microfiche, $2.00)

No. 46 *Correlates of Student Interest in Social Issues,* by R. H. Fenske, & J. F. Carmody (NAPS No. 01778; photo, $5.00, microfiche, $2.00)

No. 47 *The Impact of College on Students' Competence to Function in a Learning Society,* by M. H. Walizer, & R. E. Herriott (NAPS No. 01779; photo, $5.00; microfiche, $2.00)

No. 48 *Enrollment Projection Models for Institutional Planning,* by M. D. Orwig, P. K. Jones, & O. T. Lenning (NAPS No. 01780; photo, $5.00; microfiche, $2.00)

No. 49 *On Measuring the Vocational Interests of Women,* by N. S. Cole (NAPS No. 02071; photo, $5.00; microfiche, $1.50)

No. 50 *Stages in the Development of a Black Identity,* by W. S. Hall, R. Freedle, & W. E. Cross, Jr. (NAPS No. 02072; photo, $5.00; microfiche, $1.50)

No. 51 *Bias in Selection,* by N. S. Cole (NAPS No. 02073; photo, $5.00; microfiche, $1.50)

No. 52 *Changes in Goals, Plans, and Background Characteristics of College-Bound High School Students,* by J. F. Carmody, R. H. Fenske, & C. S. Scott (NAPS No. 02074; photo, $5.75; microfiche, $1.50)

No. 53 *Toward an Integration of Theory and Method for Criterion-Referenced Tests,* by R. K. Hambleton, & M. R. Novick (NAPS No. 02075; photo, $5.00; microfiche, $1.50)

No. 54 *College Student Migration,* by R. H. Fenske, C. S. Scott, & J. F. Carmody (NAPS No. 02215; photo, $5.00; microfiche, $1.50)

No. 55 *Predictions of Performance in Career Education,* by M. R. Novick, P. K. Jones, & N. S. Cole (NAPS No. 02216; photo, $5.00; microfiche, $1.50)

No. 56 *Predictors of Graduation from College,* by E. Nicholson (NAPS No. 02217; photo, $5.00; microfiche, $1.50)

No. 57 *Schooling and Subsequent Success: Influence of Ability, Background, and Formal Education,* by L. C. Solmon (NAPS No. 02218; photo, $5.00; microfiche, $1.50)

No. 58 *Common Fallacies about Heredity, Environment, and Human Behavior,* by A. Anastasi (NAPS No. 02220; photo, $5.00; microfiche, $1.50)

No. 59 *A Study of the College Investment Decision,* by W. W. McMahon, & A. P. Wagner (NAPS No. 02219; photo, $5.00; microfiche, $1.50)

No. 60 *Implementation of a Bayesian System for Decision Analysis in a Program of Individually Prescribed Instruction,* by R. L. Ferguson, & M. R. Novick

No. 61 *Nationwide Study of Student Career Development: Summary of Results,* by D. J. Prediger, J. D. Roth, & R. J. Noeth

No. 62 *Varieties of Accomplishment after College: Perspectives on the Meaning of Academic Talent,* by L. A. Munday, & J. C. Davis

No. 63 *Patterns of Concentration in Large Foundations' Grants to U.S. Colleges and Universities,* by R. Colvard, & A. M. Bennett

No. 64 *Vocational Choice Change Patterns of a National Sample of Community-Junior College Students,* by C. S. Scott, R. H. Fenske, & E. J. Maxey

No. 65 *Considerations and Procedures in National Norming: An Illustration Using the ACT Assessment of Career Development and ACT Career Planning Program, Grades 8-11,* by D. L. Bayless, J. W. Bergsten, L. H. Lewis, & R. J. Noeth

No. 66 *The Measurement of Economic Well-Being in Need Analysis Models,* by W. J. Goggin

No. 67 *Assessing the Career Interests of College Youth: Summary of Research and Applications,* by G. R. Hanson

No. 68 *Sex-Role Socialization and Employment Realities: Implications for Vocational Interest Measures,* by D. J. Prediger, & N. S. Cole

No. 69 *Predictive Validity of the ACT Tests at Selective Colleges,* by O. T. Lenning

No. 70 *Trends in the Academic Performance of High School and College Students,* by R. L. Ferguson, & E. J. Maxey

No. 71 *Declining Admissions Test Scores,* by L. A. Munday

No. 72 *Validity of the ACT Interest Inventory for Minority Group Members,* by R. R. Lamb

No. 73 *Teachers Are Psychologists, Too: On the Application of Psychology to Education,* by D. E. Hunt

No. 74 *Trends in the Academic Abilities, Background Characteristics, and Educational and Vocational Plans of College-Bound Students: 1970-71 to 1974-75,* by E. J. Maxey, L. M. Wimpey, R. L. Ferguson, & G. R. Hanson

No. 75 *Toward a Social Audit of Colleges: An Examination of College Student Outcomes in Terms of Admissions Information,* by L. A. Munday