

## Read Our Reports and Examine the Data A Response to Prediger and Cole

JOHN L. HOLLAND, GARY D. GOTTFREDSON  
and LINDA S. GOTTFREDSON  
*Johns Hopkins University*

The foregoing report (Prediger & Cole, in press) is neither an accurate account of the data nor of our interpretations. In addition, the report provides an incomplete, inaccurate, and simplified view of interest inventory construction, use, and associated problems: the lack of a consensual definition of sex-bias, the role of vocational theory and personal development, the practical problems in test construction and use, the evidence about inventory effects on users, the distinction between scientific assessment and social action, the criteria for ideal inventories, distinctions between sex differences and sex stereotypes or bias, the ethical issues involved in lowering the validity of an inventory in the service of social action without informing the users, the importance of a person's own aspirations and self-concept, and related matters. Interested readers should see the following reports and come to their own conclusions. The first report in the list provides a comprehensive overview.

### GENERAL REPORTS

Holland, J. L. The use and evaluation of interest inventories and simulations. In E. Diamond (Ed.), *Issues of sex bias and sex fairness in career interest measurement*. Washington, DC: U.S. Government Printing Office, in press. A comprehensive account of the problems inherent in the use and evaluation of interest inventories: the need for a generally acceptable and empirical definition of sex bias, the evidence of the actual influence of inventories on people, some recommendations and strategies for improving interest inventories, training counselors and planning career development programs, and needed research.

Holland, J. L. *Sex differences, sex-biases, and interest inventories*. Paper presented at the American Personnel and Guidance Association Convention, New Orleans, April, 1974. (Available from Center for Social

Request reprints from the authors, Educational Research Center, Johns Hopkins University, Baltimore, MD 21218.

Organization of Schools, Johns Hopkins University, Baltimore, Maryland 21218.) A critical discussion of papers by Cole, Tittle, and Vetter on alleged sex biases.

Holland, J. L. Some guidelines for reducing systematic biases in the delivery of vocational services. *Measurement and Evaluation in Guidance*, 1974, 4, 210-218. Discusses ideal vocational services, poor and biased services, potential remedies, including research strategies for defining, understanding, and reducing bias.

Holland, J. L. Some empirical and practical definitions of sex bias in interest inventories. A minority report to appear in E. Diamond (Ed.), *Issues of sex bias and sex fairness in career interest measurement*. Washington, DC: U.S. Government Printing Office, in press. Holland declined to endorse the NIE guidelines for lack of evidence and a clear definition of sex bias. Instead he proposed two empirical definitions. For example, an inventory is unbiased when its experimental effects on female and male respondents are similar and of about the same magnitude—that is, when a person acquires more vocational options, becomes more certain, or learns more about himself or herself and the world of work. . . . The principles can be extended to any area of bias by asking what difference proposed revisions of inventories, books, teacher and counselor training would make.

#### EFFECTS OF INVENTORIES

Zener, T. B., & Schnuelle, L. *An evaluation of the Self-Directed Search* (Research Report No. 124). Baltimore: Johns Hopkins University, Center for Social Organization of Schools, 1972. (ERIC Document Reproduction Service No. ED 061 458) This experimental investigation compares the effects on high school students of the *Self-Directed Search* (SDS), a self-administered vocational counseling simulation; the *Vocational Preference Inventory* (VPI); and no treatment. Boys and girls taking the SDS or VPI evaluated the instruments as moderately positive, reported feeling more satisfied with their current occupational choice, and were considering more occupational alternatives than the control group both the day after the experiment and three weeks later.

Redmond, R. E. Increasing vocational information-seeking behaviors of high school students (Doctoral dissertation, University of Maryland, College Park, 1972). *Dissertation Abstracts International*, 1973, 34, 2311A-2312A. (University Microfilms No. 73-17, 046.) The SDS increased vocational information seeking.

Nolan, J. J. The effectiveness of the Self-Directed Search compared with group counseling in promoting information-seeking behavior and realism

of vocational choice (Doctoral dissertation, University of Maryland, College Park, 1974). *Dissertation Abstracts International*, 1974, 35, 195A. (University Microfilms No. 74-16, 569.) No differences except one at the .10 level. Used only male sample of veterans.

Krivatsy, S. *Differential effects of three vocational counseling treatments*. Unpublished doctoral dissertation, University of Maryland, College Park, 1974. College counselors, the regular form of the SDS, and a special form of the SDS have similar positive effects on frequency and variety of information seeking, satisfaction with treatment, and other measures.

#### NORMS, NORMATIVE DATA, AND VALIDITY

Holland, J. L. *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice Hall, 1973. Presents the theory that led to the SDS. The development and validation of the classification is also described (pp. 71-82). The three-letter codes were derived not only from the VPI but also from the SVIB, the Kuder, the SDS, the *Dictionary of Occupational Titles*, and some job analyses by other investigators. The current SDS codes are a *compilation* of this evidence. If sex differences occurred in the interest data for a particular occupation, a code was established so that both sexes were represented as equally as possible.

Holland, J. L., & Nafziger, D. H. A note on the validity of the Self-Directed Search. *Measurement and Evaluation in Guidance*, 1975, 7, 259-262. The scales of the SDS correlate with the scales of the Kuder, the Thurstone Temperament Schedule, the Bennett Mechanical Comprehension Test, and the Minnesota Paper Form Board in predictable ways across three small samples of high school students. The results support the construct validity of the SDS as well as the constructs in Holland's theory.

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. Using a person's responses to interest items but no information on sex, the SDS has predictive validity for 1- and 3-year periods, and the introduction of sex-norms cuts correct predictions for college women in half. Base rate predictions that result in more correct predictions for women are not used because they use *only* information about a person's sex to make predictions.

Gottfredson, G. D., & Holland, J. L. Some normative self-report data on activities, competencies, occupational preferences, and ability ratings for high school and college students and employed men and women. *JSAS Catalog of Selected Documents in Psychology*, 1975, 5, 192. (Ms. No. 859) Provides normative data on SDS codes and component scores

for practitioners and researchers. Sex differences in a given area—for example, Realistic—persist regardless of item content—activities, self-ratings, competencies, or occupational preferences. Because of the differing life-long experiences of men and women in our culture, sex differences are currently expected.

Nafziger, D. H., Holland, J. L., & Gottfredson, G. D. Student-college congruency as a predictor of satisfaction. *Journal of Counseling Psychology*, 1975, 22, 132-139. Students who are congruent with their college majors according to an SDS assessment were slightly but significantly more satisfied than were incongruent students.

Hanson, G. R. *Assessing the career interests of college youth: Summary of research and applications* (ACT Research Report No. 67). Iowa City: American College Testing Program, 1975. The norms for the ACT Interest Inventory (table 33) show some contradictions between a person's raw scores and his or her transformed scores. For example, a woman with a standard score of 50 on Technical Interests has an average item response of 2.2 on a 5-point scale, where 5 means "like very much," 1 means "dislike very much," 2 means "dislike a little," and 3 means "indifferent." To take another example, if a woman says "indifferent" to every item on the Technical scale, she is told that she is at the 90th percentile. These and other examples in table 33 provide vivid illustrations of how transformed scores may mislead the user.

Cronbach, L. J. *Essentials of psychological testing* (Third edition). New York: Harper and Row, 1970. "Emphasizing norms is hard to defend. To tell a girl, 'your interest in raising a family is at the 30th percentile' suggests that her interest is low; this is quite the wrong conclusion, if 90 per cent of girls are positively interested in that responsibility. 'Your interest in washing dishes is at the 90th percentile' is equally misleading, if 99 per cent of girls hate dishwashing. At most, the facts suggest that this girl hates dishwashing less intensely than others, and is not so wildly enthusiastic about raising a family as some. . . . As a minimum the interpreter should know what percentile standing on each dimension represents the point where liking shifts to indifference" (p. 486).

#### INTEREST INVENTORY ITEMS

Gottfredson, G. D. A note on sexist wording in interest measurement. *Measurement and Evaluation in Guidance*, in press. In this experiment, changing the occupational titles of an interest inventory to make them gender neutral made no difference to high school girls. They respond as frequently to "policeman" as to "police officer."

Holland, J. L., & Gottfredson, G. D. Sex differences, item revisions, validity,

and the Self-Directed Search. *Measurement and Evaluation in Guidance*, in press. This experiment used commonly suggested items (sewing machine use, gardening, etc.) to see if these proposed items scale properly or have their highest correlations with the appropriate scale. Only two items scaled properly.

Boyd, V. S. *The linguistic structure of the Self-Directed Search: A study in sex-role stereotyping*. Unpublished doctoral dissertation, University of Maryland, College Park, 1975. College women were randomly assigned to one of two treatment conditions: 133 women completed the standard SDS, and 133 women completed a revised SDS in which all "masculine-toned terminology was neutralized." No statistically significant differences were observed in the distributions of one-, two-, or three-letter codes. No differences were observed in preferences for traditional and non-traditional occupations. On the average, respondents rated both SDS forms sexually fair, but gave a slightly higher rating to the revised form.

Rayman, J. *Sex and the single interest inventory: An empirical validation of sex balanced vocational interest inventory items*. Unpublished doctoral dissertation, University of Iowa, 1974. Rayman created an inventory of six scales containing 53 sex-balanced items from an original pool of 164 items, but he assumes that such scales will have lowered predictive validity (pp. 100-101).

#### SOCIETY

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. Representative samples of vocational aspirations and census data on employment yield similar distributions across the six Holland categories, but the categories differ greatly in popularity or frequency of employment for the two sexes and for different educational levels. When inventory results diverge greatly from either people's aspirations or the distribution of actual work, they will have lowered validity.

#### THEORETICAL MISUNDERSTANDINGS

Prediger, D. J., & Hanson, G. R. Holland's theory of careers applied to women and men: Analysis of assumptions. Unpublished manuscript, American College Testing Program, Iowa City, IA 52240. The authors examine two assumptions which are *not* in the theory, and they conduct these examinations with the use of non-SDS data.

## MORE HYPOTHESES OR MORE RESEARCH?

The evidence about interest inventories contradicts some hypotheses raised by Prediger and Cole. At the same time, the Prediger and Cole paper contains other hypotheses that only new research can resolve by comparing one kind of inventory with another, or one technique with another for one or more inventories:

1. Do inventories using sex-based norms encourage more exploratory activity than unnormed inventories?
2. Does searching a classification scheme using all permutations of a three-letter code encourage as much exploratory behavior as do sex-based norms?
3. What kind of inventory construction and reporting (norms, raw scores, sex-balanced items) predicts job satisfaction most efficiently?
4. How does the number of possible vocational options presented by an inventory compare with other considerations such as sex-based norms, sex-balanced items, etc. in increasing vocational exploration?
5. Is immediate or rapid scoring a major determinant of an inventory's influences?
6. Do occupational scales have as much exploratory effect as homogenous scales?

In a recent ACT report, Hanson (1975) also suggests numerous research projects with the ACT Interest Inventory for clarifying some issues raised by Prediger and Cole (1975). For example, does the ACT Interest Inventory "stimulate career exploration," "reinforce sex-role stereotypes," stimulate "information seeking," forecast "job satisfaction"? Evidence from such research will help us develop a clearer knowledge of inventory effects rather than continuing to rely on assumed effects because a favorite construction or reporting technique was used.

The work of Hase and Goldberg (1967) illustrates that different strategies of test construction may lead to tests with similar validity. Hase and Goldberg found that four divergent test construction strategies produced four inventories with the same overall validity. This outcome contradicts many cherished beliefs, but it moves us forward.

It seems wise to proceed with the following strategies in order to collectively create a substantial and clear knowledge about what inventories do to people and how they can help to do more for people:

1. Perform more studies of the actual influence of interest inventories.
2. Perform comparative studies of inventories developed by divergent techniques.

3. Make inventories and their interpretive materials as open as possible so that the user knows as much as possible about their scoring, possible influences, and limitations. This may maximize a user's control over the treatment.

4. Try to integrate as many virtues in an inventory as possible by inventing ways to incorporate divergent goals—maximum exploration, maximum predictive value, maximum openness.

Finally, although we are in sharp disagreement with the ACT writers, we wish to emphasize that we do not have the final answers to this complex scientific and social problem. At the same time, we object to those who would prematurely force their solution on everyone else.

## REFERENCES

- Hanson, G. R. *Assessing the interests of college youth: Summary of research and applications* (ACT Research Report No. 67). Iowa City: American College Testing Program, 1975.
- Hase, H. D., & Goldberg, L. R. Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin*, 1967, 67, 231-248.
- Prediger, D. J., & Cole, N. S. *Sex-role socialization and employment realities: Implications for vocational interest measures* (ACT Research Report No. 68). Iowa City: American College Testing Program, in press.

Received: March 31, 1975.