# Dependence of the Relative Productivity Gains of Two Personnel Selection Tests on the Applicant Pool Size

Louis M. Hsu
Fairleigh Dickinson University

Schmidt, Hunter, McKenzie, and Muldrow (1979) have recently demonstrated how the use of a new test, which differed from a previous test in terms of validity and/or per applicant cost, could result in impressive gains in productivity (utility). This paper focuses on the consequences of changing the applicant pool size (keeping the number of selectees fixed) on the relative productivity gains of the two tests. It is shown that the utility gain may be larger for one test than for the other for part of the range of possible applicant pool sizes and smaller for the rest of that range. Methods are described for determining for any two tests (1) whether such a reversal can occur and (2) the range of applicant pool sizes leading to greater utility gains for each test over the other. An implication is that the choice of a test should be contingent on an analysis of the relative productivity gains of the competing procedures for the available applicant pool sizes.

Chen and Novick (1982) recently noted that "decision theory can make important contributions to . . . employment decision making, provided one can quantify the utilities of different possible outcomes" (p. 19). The definition and determination of utility functions has been viewed as the "major problem in using utility theory" (Chen & Novick, 1982). Many authors have addressed themselves to this problem. Recently Cronbach, Yallow, and Schaeffer (1980), Gross and Su (1975), Petersen (1976), Huynh (1976, 1977), Petersen and Novick (1976), and Novick and Petersen (1976) have proposed "threshold utility" models, in which utility is viewed as a two-valued function of criterion performance. Novick and Lindley (1978, 1979) have described what they consider to be "more realistic forms that utility functions might take in selected . . . applications" (p. 181). Their model defined utility for members of a group in terms of a normal ogive whose parameters are specified by decision makers (administrators) for that group. They outlined advantages of their normal-distribution-based utility function over the "threshold utility" functions.

## The Brogden-Cronbach-Gleser Model

Another model in which the utility function is associated with a normal distribution function was developed by Cronbach and Gleser (1965) from an earlier model of Brodgen (1946, 1949). This model allows estimation of the utility, in dollars, of personnel selection procedures. Recently the Brogden-Cronbach-Gleser model has received much attention (see Bobko, Karren, & Parkinson, 1983; Cascio,

359

1980; Cascio & Silbey, 1979; Holgarth & Einhorn, 1976, Hunter, Schmidt, & Rauschenberger, 1977; Landy, Farr, & Jacobs, 1982; Schmidt, Hunter, & Pearlman, 1982). The resurgence of interest in this utility model (which measures utility in dollars) is perhaps due to the "decline in the rate of economic growth in the United States" (Bobko et al., 1983) and the consequent importance to both organizations and consumers of economically productive personnel selection strategies (see Landy et al., 1982).

It should be noted that most authors who currently discuss and use the Brogden-Cronbach-Gleser model do not distinguish between the monetary value (or expected monetary value) and the utility (or expected utility) associated with the use of personnel selection tests. In fact, they generally identify utility with monetary value in most of their writings. Although monetary gains (as well as criterion or job performance) and utility have been clearly distinguished for some time (see Schlaifer, 1959) and although this distinction is evident in the work of Chen and Novick (1982), Chuang, Chen, & Novick (1981), Gross and Su (1975), Petersen (1976), Huynh (1976, 1977), Petersen and Novick (1976), and Novick and Petersen (1976), it is generally recognized that there are circumstances under which these concepts may be identified or viewed as linearly related (see Cronbach et al., 1980; Savage, 1968, p. 27; Schlaifer, 1959, pp. 28–31). In any case, it should be noted that the term "utility," in the context of the Brogden-Cronbach-Gleser model and most current discussions and applications of this model (Bobko et al., 1983; Cascio, 1980; Cascio & Silbey, 1979; Holgarth & Einhorn, 1976; Hunter et al., 1977; Landy et al., 1982; Schmidt et al., 1982), refers to "monetary," "productivity," "economic," or "job performance" variables or to linear functions of these variables.

A major obstacle to the use of the utility functions of Brogden and of Cronbach and Gleser has been the estimation of the parameter $SD_y$, the standard deviation of job performance measured in dollars (Dunette & Borman, 1979, p. 493; Kaplan & Saccuzzo, 1982, p. 166). Early approaches to the estimation of this parameter (Roche, 1965) involved complex and difficult-to-implement procedures (Schmidt, Hunter, McKenzie, & Muldrow, 1979). Recently, however, Schmidt et al. (1979), Schmidt et al. (1982), and Hunter and Schmidt (in press) have developed useful and efficient methods of estimating this parameter. Even more recently Bobko et al. (1983) have provided evidence of the tenability of assumptions underlying some of Schmidt and coworkers' methods of estimating this standard deviation. Using one of these methods, Schmidt et al. (1979) showed that large gains in utility could result from using a new instead of a previous personnel selection test, given fixed validity coefficients and costs associated with these tests.

### Purpose

The purpose of this article is to investigate how the difference in the utility gains of two personnel selection (or classification) tests is affected by changes in the applicant pool size. A situation in which this topic is of practical interest occurs when a fixed number of applicants is to be selected but the number of persons screened may vary (Cronbach & Gleser, 1965, pp. 40–41; Hogarth & Einhorn, 1976) within a certain range. The variable number of screened applicants may result from (1) availability of a large number of persons within an institution (e.g., a branch of the armed services or of the government, or a large corporation); (2) changes in recruitment efforts (Hogarth & Einhorn, 1976); or (3) the influence of many factors determining the number of applicants who will apply for a job. Limits on the sizes of the pools of applicants may be determined by such factors as the existing number of trained personnel in a geographical location (Kaplan & Saccuzzo, 1982, p. 167).

The assumptions underlying the derivation of Equation 6 of Schmidt et al. (1979) and the derivation and application of Equation 1.13 of Cronbach and Gleser (1965, p. 309, pp. 40–41) will be considered tenable in the present paper. One additional assumption is that the cost per screened applicant (as opposed to the cost per selectee) is not a function of the applicant pool size. Implications of a less restrictive

assumption will be considered in the Discussion. It will be shown that the less restrictive assumption leads to the same conclusions.

## Determination of Applicant Pool Sizes for Which the Two Tests Yield Equal Utility Gains

Cronbach and Gleser (1965, p. 309) note that the gain in utility per selectee, given test $i$ is (using Schmidt et al.'s, 1979, p. 614, notation)

$$\Delta \overline{U}/\text{selectee} = (SD_y \ r_i \ (\phi/p) \ - \ C_i/p) \qquad [1]$$

where

$SD_y$ = standard deviation of job performance, measured in dollars;
$r_i$ = correlation coefficient (validity) of test $i$ with $y$;
$\phi$ = ordinate of the standard normal distribution corresponding to the selection ratio;
$p$ = selection ratio; and
$C_i$ = average cost associated with screening an applicant with test $i$ (see Van Naerssen, 1965, pp. 282–284, for illustration of some factors included in this cost).

Therefore, the difference in utility gains (per selectee) of Tests 1 and 2 would be

$$D = (SD_y r_1 (\phi/p) \ - \ C_1/p) \ - \ (SD_y r_2 (\phi/p) \ - \ C_2/p) \qquad [2]$$

Thus, the difference in utility gains for Tests 1 and 2 can be zero, provided Equation 2 can equal zero. To determine when this can happen, Equation 2 may be solved for $\phi$, when $D = 0$, designating the result $\phi_o$:

$$\phi_o = (C_1 \ - \ C_2)/(SD_y (r_1 \ - \ r_2)) \qquad [3]$$

Since $\phi$ in Equation 1 is the ordinate of the standard normal distribution, and since the range of possible values of $\phi$ is therefore

$$0 \ \le \ \phi \ \le \ 0.3989 \qquad [4]$$

it is clear that Tests 1 and 2 can be of equal utility (barring the trivial case of $\phi = 0$) only if

$$0 \ < \ \phi_o \le \ 0.3989 \qquad [5]$$

Given that $\phi_o$ is within this range, the selection ratios for which the two tests will be of equal utility may be determined from standard normal tables: these selection ratios may be designated $p_o$ and $(1 - p_o)$.

If a fixed number of applicants ($N_a$) is to be selected, the applicant pool sizes for which the tests will be of equal utility will therefore be

$$(1/p_o)N_a \quad \text{and} \quad (1/(1 \ - \ p_o))N_a \qquad [6]$$

## Determination of Applicant Pools Sizes for Which the Utility of Each Test Exceeds That of the Other

It is clear from Equation 2 that Test 1 will yield greater utility gain than Test 2 whenever $D > 0$, and smaller utility whenever $D < 0$. Rewriting Equation 2 as

$$(1/p) ( \phi (r_1 \ - \ r_2) \ SD_y \ - \ (C_1 \ - \ C_2)) = D, \qquad [7]$$

and noting that $p$ is always positive, it is then apparent that Test 1 will yield greater utility gain than Test 2 when

$$(\phi (r_1 - r_2) \, SD_y - (C_1 - C_2)) > 0 \qquad [8]$$

that is, if and only if

$$\phi > \left( \frac{(C_1 - C_2)}{SD_y \, (r_1 - r_2)} \right) = \phi_0 \qquad [9]$$

Similar reasoning indicates that Test 2 will yield greater utility gain than Test 1 if and only if,

$$\phi < \left( \frac{(C_1 - C_2)}{SD_y \, (r_1 - r_2)} \right) = \phi_0 \qquad [10]$$

Equation 9 implies that Test 1 will have greater utility than Test 2 for selection ratios associated with $\phi > \phi_0$, i.e., for selection ratios from

$$p_0 \quad \text{to} \quad (1 - p_0). \qquad [11]$$

Similarly, Equation 10 implies that Test 2 will have greater utility gain than Test 1 for selection ratios outside the range

$$p_0 \quad \text{to} \quad (1 - p_0). \qquad [12]$$

Thus, the applicant pool sizes (*N*s) for which Test 1 will yield greater utility gain than Test 2 will be in the interval

$$(1/p_0) \, N_a \quad > \quad N \quad > \quad (1 / (1 - p_0)) \, N_a \qquad [13]$$

(where $N_a$, it may be recalled, is the number of selectees). Similar reasoning shows that the applicant pool sizes for which Test 2 has greater utility gain than Test 1 will consist of

$$(N > (1/p_0) \, N_a) \quad \text{and} \quad (N < (1 / (1 - p_0)) \, N_a) \qquad [14]$$

In summary, (1) if $\phi_0$ (obtained using Equation 3) is not within the range defined in Equation 4, then one test will yield greater utility gains per selectee than the other, regardless of the applicant pool size; and (2) if $\phi_0$ is within the range defined in Equation 4, then (a) Equation 13 may be used to determine the range of applicant pool sizes for which Test 1 will yield greater utility gains than Test 2 and (b) Equation 14 may be used to determine the range of applicant pool sizes for which Test 2 will yield greater utility gains than Test 1.

### Illustration

Consider two competing tests: The validity coefficient of Test 1 as a predictor of job performance is $r_1 = .76$, and for Test 2, $r_2 = .5$ (cf. values of *r*s in Schmidt et al.'s, 1979, p. 622, examples). Assume that the standard deviation of job performance is $10,000 (cf. Schmidt et al., 1979, p. 622) and that the costs associated with the two assessment methods are, respectively, $1,444 and $700 per assessee (cf. Schmidt et al., 1979, example on p. 614).

Equation 3 indicates that $\phi_o = [(1444 - 700)/(10000(.76 - .5))] = .2862$ if the two tests have equal utility. Since $\phi_o$ is in the range $[0 < \phi < .3989]$, there will be applicant pool sizes for which the two tests will have equal utility gains. Standard normal distribution tables indicate that the selection ratios corresponding to an ordinate of .2862 are $p_o = .2076$ and $(1 - p_o) = .7924$. Thus if $N_a = 400$ applicants are to be hired, the two tests will be of equal utility if the applicant pool sizes are (applying Equation 6) $(1/.2076)(400) \cong 1927$, or $(1/.7924)(400) \cong 505$. Equation 13 indicates that Test 1 will have greater utility gain than Test 2 if the applicant pool size is between 505 and 1,927, and Equation 14 shows that Test 2 will yield greater utility gain than Test 1 if the applicant pool size is less than 505 or greater than 1,927.

Figure 1 illustrates in detail the relation of utility gain of Test 1 to utility gain of Test 2, as the applicant pool size changes. Figure 1 confirms results obtained with Equations 3, 13, and 14; that is, Figure 1 shows that (1) the tests yield equal utility gains when the applicant pool size is 505 or 1,927, (2) Test 1 yields greater utility gain than Test 2 for applicant pool sizes between 505 and 1,927, and (3) Test 2 yields greater utility gain than Test 1 for applicant pool sizes outside that range.

## Discussion

It is clear, from examination of Equation 7, that if the costs associated with the two tests *or* if the validity coefficients of these tests are equal (but not both), then one test will yield consistently greater

## Figure 1
### Relation of Utility Gain per Selectee to Applicant Pool Size
### for Two Combinations of Validity Coefficients and Costs per Applicant
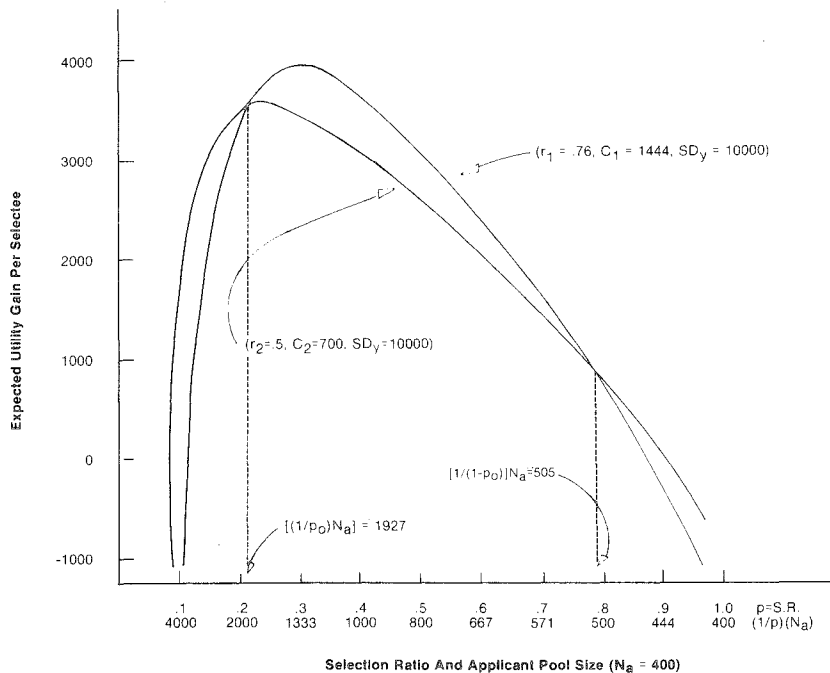
utility gains than the other, regardless of the applicant pool size ($N$). However, when neither the costs, nor the validity coefficients of the tests are equal, reversals in the relative sizes of utility gains of such tests may occur with changes in the applicant pool size; that is, the utility gain of Test 1 may exceed that of Test 2 for some applicant pool sizes, whereas the utility gain of Test 2 may exceed that of Test 1 for other applicant pool sizes.

Determination of whether reversals could occur given two competing tests (of fixed validities and associated costs per applicant) can be made using Equation 3, since the occurrence of reversals implies that the utility gain curves of the two tests (see Figure 1) will cross. If reversals occur for selection ratios $> 0$ but $< 1$, then $\phi_0$ (obtained by Equation 3) will be within the range $[0 < \phi_0 < .3989]$. Should $\phi_0$ not be within this range, no reversals will occur, and a decision about which test leads to greater utility gains can be based on results of *any* fixed selection ratio study of the utility gains of the two tests. However, if $\phi_0$ is between 0 and .3989, reversals will generally occur. Then the choice of test depends on whether (1) the applicant pool size can be controlled at will (as might be the case in the armed services); (2) certain constraints exist on applicant pool sizes, but selection may be from among several possible pool sizes (as might be the case if there could be selection from several geographical locations differing in availability of potential applicants); or (3) the applicant pool size can only be estimated but not controlled (as might be the case when the effects of a fixed job advertisement budget could be estimated from past experience).

In Case 1 the test with the larger maximum utility gain should be used (other things being equal) and the applicant pool size should be selected to yield the maximum utility gain. Cronbach and Gleser (1965, p. 309, Equation 1.13) have shown that the optimal selection ratio for test $i$ with validity $r_i$ and cost per applicant $C_i$ will occur when

$$\phi - z p = (C_i / (SD_y r_i)). \qquad [15]$$

where $z$ is the $z$-normal score corresponding to the selection ratio. Van Naerssen (1965) provided a table (p. 275) for determining $p$ values from knowledge of the right-hand term of Equation 15. Application of Equation 15 and use of Van Naerssen's table yields the applicant pool sizes (determined from values of $p_i$) which maximize the utility gains of the two tests ($N_{max_i}$). These applicant pool sizes will generally be different for the two tests. Substitution of the $N_{max_i}$ values in Equation 1 (note that $p = N_a/N$) determines which of these tests yields the greater maximum utility gain. The test with the greater gain should be used, with the applicant pool size yielding that maximum.

In Case 2, utility gain curves (such as those in Figure 1) could be generated using Equation 1 to determine which combination of test and applicant pool size yields the maximum utility gain, given the available applicant pool sizes. In Case 3 the anticipated applicant pool size could be compared to the ranges defined by Equations 13 and 14. If the predicted applicant pool size is within the range defined by Equation 13, then Test 1 could be expected to yield the greater utility gain and therefore should be used; if it is within the range defined in Equation 14, then Test 2 would have the higher utility gain and should therefore be used.

An assumption which was made in deriving equations in this article was that $C_i$, the average cost associated with testing an applicant with test $i$, was not a function of the applicant pool size. However, exactly the same equations could be derived and used with the less restrictive assumption that only $(C_1 - C_2)$ was not a function of the applicant pool size, even though both $C_1$ and $C_2$ could be functions of the applicant pool size. All comments in the discussion are thus applicable to this more general situation. If it cannot be assumed that $(C_1 - C_2)$ is constant across applicant pool sizes, but if $C_1$ and $C_2$ are known functions of applicant pool sizes, test selection could still be performed after generation of the utility gain curves of the two tests.

# References

Bobko, P., Karren, R., & Parkinson, J. J. Estimation of standard deviations in utility analyses: An empirical test. *Journal of Applied Psychology*, 1983, *68*, 170–176.

Brogden, H. E. On the interpretation of the correlation coefficient as a measure of predictive efficiency. *Journal of Educational Psychology*, 1946, *37*, 65–76.

Brogden, H. E. When testing pays off. *Personnel Psychology*, 1949, *2*, 171–183.

Cascio, W. F. Responding to the demand for accountability: A critical analysis of three utility models. *Organizational Behavior and Human Performance*, 1980, *25*, 32–45.

Cascio, W. F., & Silbey, V., Utility of the assessment center as a selection device. *Journal of Applied Psychology*, 1979, *64*, 107–118.

Chen, J. J., & Novick, M. R. On the use of a cumulative distribution as a utility function in educational or employment selection. *Journal of Educational Statistics*, 1982, *7*, 19–35.

Chuang, D. T., Chen, J. J., & Novick, M. R. Theory and practice for the use of cut scores for personnel decisions. *Journal of Educational Statistics*, 1981, *6*, 129–152.

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

Cronbach, L. J., Yallow, E., & Schaeffer, G. A mathematical structure for analyzing fairness in selection. *Personnel Psychology*, 1980, *33*, 693–704.

Dunnette, M. D., & Borman, W. C. Personnel selection and classification systems. *Annual Review of Psychology*, 1979, *30*, 489–522.

Gross, A. L., & Su, W. H. Defining a "fair" or "unbiased" selection model: A question of utilities. *Journal of Applied Psychology*, 1975, *60*, 345–351.

Hogarth, R. M., & Einhorn, H. J. Optimal strategies for personnel selection when candidates can reject offers. *Journal of Business*, 1976, *49*, 478–495.

Hunter, J. E., & Schmidt, F. L. Fitting people to jobs: The impact of personnel selection on national productivity. In E. A. Fleishman (Ed.), *Human performance and productivity*. Hillsdale NJ: Erlbaum, in press.

Hunter, J. E., Schmidt, F. L., & Rauschenberger, J. M. Fairness of psychological tests: Implications of four definitions for selection utility and minority hiring. *Journal of Applied Psychology*, 1977, *62*, 245–260.

Huynh, H. Statistical consideration of mastery scores. *Psychometrika*, 1976, *41*, 65–78.

Huynh, H. Two simple classes of mastery scores based on the beta-binomial model. *Psychometrika*, 1977, *42*, 601–608.

Kaplan, R. M., & Saccuzzo, D. P. *Psychological testing*. Monterey CA: Brooks/Cole, 1982.

Landy, F. J., Farr, J. L., & Jacobs, R. R. Utility concepts in performance measurement. *Organizational Behavior and Human Performance*, 1982, *30*, 15–40.

Novick, M. R., & Lindley, D. V. The use of more realistic utility functions in educational applications. *Journal of Educational Measurement*, 1978, *15*, 181–192.

Novick, M. R., & Lindley, D. V. Fixed-state assessment of utility functions. *Journal of the American Statistical Association*, 1979, *74*, 306–311.

Novick, M. R., & Petersen, N. S. Towards equalizing educational and employment opportunity. *Journal of Educational Measurement*, 1976, *13*, 77–88.

Petersen, N. S. An expected utility model for "optimal" selection. *Journal of Educational Statistics*, 1976, *4*, 333–358.

Petersen, N. S., & Novick, M. R. An evaluation of some models for culture-fair selection. *Journal of Educational Measurement*, 1976, *13*, 3–31.

Roche, W. J. A dollar criterion in fixed-treatment employee selection. In L. J. Cronbach & G. C. Gleser (Eds.), *Psychological tests and personnel decisions*. Urbana: University of Illinois Press, 1965.

Savage, I. R. *Statistics: Uncertainty and behavior*. New York: Houghton Mifflin, 1968.

Schlaifer, R. *Probability and statistics for business decisions: An introduction to managerial economics under uncertainty*. New York: McGraw-Hill, 1959.

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. Impact of valid selection procedures on work force productivity. *Journal of Applied Psychology*, 1979, *64*, 609–626.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. Assessing the economic impact of personnel programs on work force productivity. *Personnel Psychology*, 1982, *35*, 333–347.

Van Naerssen, R. F. Application of the decision-theoretical approach to the selection of drivers. In L. J. Cronbach & G. C. Gleser, (Eds.), *Psychological tests and personnel decisions*. Urbana: University of Illinois Press, 1965.

## Acknowledgements

## Author's Address

Send requests for reprints or further information to Louis M. Hsu, Fairleigh Dickinson University, Teaneck NJ 07666, U.S.A.