PERSONNEL PSYCHOLOGY
1993, 46

EXPANDING THE CONTEXT OF UTILITY: THE STRATEGIC IMPACT OF PERSONNEL SELECTION

CRAIG J. RUSSELL
Department of Management
Louisiana State University

ADRIENNE COLELLA, PHILIP BOBKO
Department of Management
Rutgers University

Others have demonstrated that traditional applications of the Brogden-Cronbach-Gleser (BCG) selection utility formula are deficient in responding to the financial context of managerial decisions (Boudreau, 1983a, 1983b; Cronshaw & Alexander, 1985, 1991). We demonstrate that traditional estimates of selection utility also fail to reflect the strategic context faced by managerial decision makers. We modify the traditional BCG model to yield an estimate of total utility derived from human resources (U_{total}) that can be directly compared to firms' strategic need at a particular point in time (U_{target}). Further, we demonstrate that, while strategic need is rarely constant over time, the capacity of a selection system to meet that need is also likely to change as r_{xy} and SD_y change over time. Re-examination of what is important to strategic human resource decision makers (selection utility vs. total utility and strategic need) and changing selection system contributions over time yields a more realistic view of how firms benefit from personnel selection.

Personnel researchers have been interested in assessing the financial impact of selection methods since the 1920s. Most activity from 1920–1965 focused on developing models of utility (e.g., Brogden, 1946; Cronbach & Gleser, 1965). These models typically portrayed a relationship between two sets of constructs—psychological "predictor" constructs and economic "criterion" constructs reflecting costs and benefits. Unfortunately, much as "criterion problems" have plagued the larger

A portion of this study was funded by the Navy Personnel Research and Development Center under Battelle contract #DAALO3-86-D-0001. The opinions expressed in this article are those of the authors, are not official, and do not necessarily reflect the views of the Navy Department. We would like to thank Arthur Bedeian, Jack Feldman, Kevin Mossholder, Dirk Steiner, and three anonymous reviewers for their comments on an earlier draft. The authors made equal contributions to this study; order of authorship is reverse alphabetical.

Correspondence and requests for reprints should be addressed to Craig Russell, Department of Management, College of Business, Louisiana State University, Baton Rouge, LA 70803-6312; Adrienne Colella and Philip Bobko can be reached at the Department of Management, School of Business, Rutgers University, New Brunswick, NJ 08903.

domain of personnel selection (Smith, 1976), problems in identifying, operationalizing, and incorporating aspects of the economic criterion construct domain have plagued efforts at estimating selection utility. For example, Boudreau (1991) and Cronshaw and Alexander (1985) argued that current personnel selection utility models are deficient when they ignore the decision context in which utility estimates are used. Boudreau (1983a) and Cronshaw and Alexander (1985) specifically introduced capital budgeting techniques from the field of financial planning to correct part of this deficiency. Boudreau (1983b) made modifications to incorporate the likelihood of changes in labor flow over time, while Cronshaw, Alexander, Wiesner, and Barrick (1987) and Rich and Boudreau (1987) demonstrated alternate means of assessing risk or variability in the point estimates of selection utility.

In summarizing this research Boudreau (1991) noted that when utility models "focus on the consequences of improving the... labor force, a fundamental consideration is how the organization uses quality improvements" and that "all existing payoff scales reflect a concern with productivity-based outcomes, virtually ignoring other factors that might be affected by selection decisions, such as community relations, work force attitudes, and adherence to a code of ethics. Thus, every payoff function is deficient in some way" (p. 648). We would add that no current models of selection utility account for the strategic needs of the firm. Specifically, while certain generic economic objectives are common to all private sector organizations (e.g., profit maximization, cost minimization), strategic opportunities are not common across firms and do not occur within firms in a uniform, predictable way (Ansoff, 1988). As opportunities are forecasted to occur, strategic decision makers must carefully align resources (personnel, capital, and raw materials) to meet strategic objectives. As strategic objectives vary within and between firms, different "alignments" of resources are dictated.

For example, modifying an illustration from Ansoff (1988), an entrepreneur starting a mail-order computer firm may see a strategic opportunity to develop a distribution network, gaining enough market share to meet the strategic objective of selling out to a major computer manufacturer within a 3-year period while reaping a substantial capital gain. It would be expected that the way this entrepreneur "aligns" resources, including the recruiting and selection of critical personnel, would be very different from the way resources would be "aligned" if the strategic objective was to grow large enough to pose a long-term challenge to major computer manufacturers. The entrepreneur concerned with capital gains would need a narrow applicant pool (i.e., one with a fairly high minimum skill level) and a selection system that procured skilled personnel to meet needed performance levels relatively quickly (at least within a 3-year period). Another entrepreneur concerned with long-term growth could focus on a broader applicant pool with an emphasis on hiring less skilled applicants that are trained and developed over a longer period. Traditional applications of utility models to the selection of computer sales personnel will be deficient because they fail to account for differences in the entrepreneurs' strategic objectives.

The human resource literature has recently begun to address the strategic implications of human resource systems (Devanna, Frombrun, & Tichy, 1981; Schuler & Jackson, 1987; Snow & Snell, 1993; Tsui & Gomez-Mejia, 1988). However, integrations of strategic considerations and methods of estimating selection utility have not been considered. It is very likely that estimates of utility will be identical for a given selection system in the two strategic scenarios outlined above (given identical predictors and criteria), especially if estimates of criterion validity have been meta-analytically derived and corrected for range restriction caused by the "narrow" versus "broad" applicant pools found in the two strategic circumstances. Very simply, a selection system that did not result in performance levels needed to attract buy-out offers would be worthless to our entrepreneurs concerned with capital gains, while potentially very valuable to entrepreneurs concerned with long-term growth potential.

Ansoff (1988) elegantly described how changes in size, technology, and third party ownership of firms have made profit maximization only one of many strategic objectives. Ansoff (1988, pp. 31–32) provided a taxonomy of organizational objectives that might result. Different strategic intent (Hamel & Prahalad, 1989), such as the near-term buy out versus long-term growth objectives of our computer entrepreneurs, will embrace different weighted combinations of these goals and objectives.

The primary purpose of the current effort is to demonstrate how a firm's strategic considerations change the focus of personnel selection utility discussions. Derivations resulting from this change provide strategic decision makers with additional information needed to determine a selection system's contribution to strategic goals and objectives. First, we introduce the notion of organizations' strategic goals, how strategic goals might vary over time, and how strategic goals impact the application of selection utility models. This change in focus captures the fact that *temporal changes* in strategic goals must be considered in determining whether a selection system contributes to the "alignment" of resources needed to meet that need.

Second, we demonstrate how existing utility models might be applied in ways that reflect the context of strategic intent. Specifically, we show how a given selection system may vary in its ability to meet changing strategic needs over time. This highlights an assumption made by current utility models that is rarely questioned: That predictor-criterion relationships (r_{xy}) and applicant pool performance variation (SD_y) can be counted on to remain the same over time (see Henry & Hulin, 1987, for an exception). Since the timing of resource "alignment" is critical if strategic needs are to be met, differences in criterion validity (r_{xy}) and applicant pool performance variance (SD_y) over time must be considered. We briefly explore the literature suggesting that changes in validity (r_{xy}) and performance variation (SD_y) might be expected over time, before illustrating the impact of such changes on utility estimates. The relaxation of this assumption in the dominant Brogden-Cronbach-Gleser utility model, combined with a focus on the strategic impact of selection utility, yields more realistic portrayals of a selection system's impact on organizational goals and objectives.

In this way, we demonstrate how joint changes in strategic need, and selection systems' capacity to meet that need over time, must be considered by human resource and strategic decision makers. Current metrics of selection utility, pegged to a lone performance measure taken at a single point in time, fail to provide decision makers with this strategic information. Before introducing these proposed changes, we briefly review the dominant Brogden-Cronbach-Gleser model of selection utility.

The Brogden-Cronbach-Gleser Model in a Strategic Context

Brogden (1946, 1949) developed a basic formula for deriving the economic return realized from a selection system:

$$\overline{Y}_s = r_{xy} SD_y \overline{Z}_{x_s} + \mu_y \tag{1}$$

where $\overline{Y}_s$ is the average job performance measured in dollar terms for the individuals actually selected, r_{xy} is the Pearson product moment correlation between the predictor X and criterion Y, SD_y is the standard deviation of the dollar performance criterion Y that would be expected if the entire applicant cohort had been hired and placed on the job, $\overline{Z}_{x_s}$ is the average predictor score performance in standard score form of those selected, and μ_y is the expected level of job performance in dollars for a group of randomly selected applicants. Note that the average individual's job performance is neatly partitioned into the portion that would have been expected if random selection had been used (μ_y) and the portion by which this "base line" performance is expected to increase due to use of the selection system ($r_{xy} SD_y \overline{Z}_{x_s}$). To obtain "net performance" or utility gain due to use of the selection system for a given time period:

$$\overline{Y}_s - \mu_y = r_{xy} SD_y \overline{Z}_{x_s}$$

or, more commonly,

$$\Delta U_{per\ selectee} = r_{xy} SD_y \frac{\lambda}{\phi} \tag{2}$$

where,

$\Delta U_{per\ selectee}$ = expected economic return relative to random selection
λ = height of normal curve at cut score
ϕ = the selection ratio
$\frac{\lambda}{\phi} = \overline{Z}_{x_s}$ = average predictor score performance in standard score form

To determine the total marginal utility resulting from use of the selection device, Brogden (1949) multiplied Equation 2 by the number of employees selected (N_s) and subtracted testing costs as follows:

$$\Delta U_{selection} = N_s SD_y r_{xy} \frac{\lambda}{\phi} - N_{app} C_s \tag{3}$$

where C_s is the cost of testing an applicant, N_{app} is the number of applicants, and $\Delta U_{selection}$ is the difference between the "total" performance level of all those selected ($N_s \overline{Y}_s$) minus the value of the base-line performance level that would have been expected if the selection system had not been used (i.e., $N_s \mu_y$, the total performance level expected from random selection). Cronbach and Gleser (1965) extended Brogden's (1946) model to other types of selection situations (e.g., multi-stage selection and placement decisions), deriving a formula identical to Equation 3. This equation has become known as the Brogden-Cronbach-Gleser (BCG) utility model, which constitutes the foundation for most utility analyses performed today (see Cascio, 1991, for the BCG model chronology).

It is interesting to note that Equation 1 can be modified in a similar manner to obtain the total utility of labor to the firm for a particular job:

$$U_{total} = [N_s r_{xy} SD_y \overline{Z}_{x_s} - N_{app} C_s] + N_{employed}\ [\mu_y - C_{labor}] \tag{4}$$

where $N_{employed}$ is the total number of individuals employed in the job and C_{labor} is the average employee cost of labor excluding the cost of the selection system (salaries, benefits, recruiting, human resource department overhead, etc.). Where $\overline{Y}_s$ in Equation 1 represents the total average performance expected from those selected, U_{total} in Equation 4 represents the *net* total performance after adjusting for the costs of that performance (i.e., C_s is the cost of the performance gain expected from use of the selection system while C_{labor} is the cost of performance levels expected from random selection). Hence, this equation is partitioned to illustrate the net value added by the selection system $(N_s r_{xy} SD_y \overline{Z}_{x_s} - N_{app} C)$ and the base-line utility expected from labor selected at random $(N_{employed}[\mu_y - C_{labor}])$. A new cohort of employees screened by the selection system and hired into a job is expected to add $(r_{xy} SD_y \overline{Z}_x - C) + (\mu_y - C_{labor})$ per employee to the firm, while all initial employees who had not been screened would each contribute only $(\mu_y - C_{labor})$. This model is similar to one developed by Boudreau and Rynes (1985) and Boudreau and Berger (1985) to estimate the utility of recruiting policies (see Boudreau & Rynes, 1985, Equation 2, p. 358), though for clarity we have deleted modifications for net present value, taxes, and changes in employee flow. Regardless, an interesting implication is that $\Delta U_{selection}$ can be large while U_{total} might be very small or even negative, a circumstance that becomes important when we introduce firms' strategic needs below.

Investigators have made various adjustments to the BCG model (Equation 3) to make it less deficient and more representative of organizational decision contexts (Boudreau, 1991). For example, Schmidt, Hunter, McKenzie, and Muldrow (1979) were the first to expand the scope of the BCG model over time by multiplying utility gains by the average tenure of employees hired under a new selection system. A key assumption (which we will come back to) of Schmidt et al. (1979) was that both criterion validity (r_{xy}) and SD_y remain constant over time, and hence, utility (Equation 3) realized in each work period by those selected remains constant over the planning period for which the utility estimate is derived. Murphy (1986) adjusted the BCG model to incorporate job offer rejections. Boudreau (1983a) and Cronshaw and Alexander (1985, 1991) argued that the BCG model should be modified to reflect capital budgeting variables, in effect enhancing the construct validity of all the BCG model's economic parameters. Boudreau (1983a) introduced the notions of variable costs (increased productivity may mean higher wages due to merit pay, lower raw material costs due to decreased waste, etc.) and taxes, while Boudreau (1983b) and Cronshaw and Alexander (1985, 1991) suggested that forecasts of future utility gained from selection of current employees be discounted. These extensions emphasize ratios and metrics commonly used by financial planners (Boudreau, 1991). This paper extends this line of inquiry by showing that inattention to the strategic context of selection decisions is an additional deficiency of the BCG model.

Change in Focus:
Strategic Need and Temporal Change in Strategic Need

A common deficiency in estimating the impact of selection systems occurs when estimates derived from the BCG model (Equation 3) are appraised without considering the expected total value added by human resources to be realized by the firm (Equation 4). As noted above, situations can occur at different points in time where the relative BCG utility estimate is positive while the absolute economic return is negative. The three examples below demonstrate this in circumstances where three different strategic objectives are dominant: economic survival, maintenance of the status quo, and rapid growth of market share through the sale of franchises.

The Ultimate Strategic Need: Economic Survival

Russell and Domm (1990) reported a criterion-related validity coefficient of .32 between an assessment center rating and profit generated by a sample of retail store managers. Let us assume that all fixed and variable costs of producing sales volume are the same across stores (Russell & Domm controlled for store size in deriving their validity coefficient). Then it can be assumed that profit generated by the store is a reasonable measure of value added by the manager. Russell and Domm reported that the validity coefficient of .32 translated into an expected quarterly increase ($\Delta U_{selectee}$) of \$3,000 in store profits. Interestingly, average quarterly profit generated by the sample was *negative* \$3,662.61 and the gain expected from use of the assessment center was not enough to offset the loss each store was already incurring. It is important to realize that these figures yield a large positive BCG utility estimate from Equation 3, though total value added by human resources to the firm (Equation 4) remains negative. Hence, relying only on the incremental utility obtained from selection might cause one to recommend investing in a costly assessment center for stores that will continue to lose money.

Most corporate strategic goals and objectives involve something beyond mere economic survival (e.g., the near-term and long-term goals of our mail-order computer entrepreneurs described above). Based on careful forecasts of environmental changes and resource availability, most firms' objectives will consist of a longitudinal pattern of desired economic outcomes, growth, product mix, and so forth. We will label this desired mix of performance levels the firm's *strategic need* over some relevant time horizon. This suggests the possibility that temporal changes in strategic need might dictate differential weighting (across time periods) of expected future economic returns from a selection system. Specific examples are described below of how differential weights applied to BCG forecasts of economic return yield very different conclusions. Each example assumes that the appropriate values of r_{xy} and SD_y have been used in each time period to derive forecasts of $\Delta U_{selection}$ (Equation 3) and U_{total} (Equation 4).

Turnover and Maintenance of Status Quo

In this scenario, over the next 2 years an organization faces high turnover due to the retirement of a disproportionately large number of its retail store managers. We will consider this firm to be focused on a narrow product market, existing in a highly homogeneous external environment, with few adjustments needed in technology, methods, or structure. Primary attention is paid to maximizing the efficiency of existing operations (i.e., maintaining the status quo). Because of this strategic posture, all future economic returns from a selection system will be of equal strategic value—the critical, or minimally acceptable level of strategically needed U_{total}, hereafter labeled U_{target}, is constant over time. These conditions are presented graphically in Figure 1.

A number of characteristics of Figure 1 are of interest. First, the forecasted utility using BCG Equation 3 is positive ($\Delta U_{selection} = \overline{Y}_s - \mu_y$) at every point in time, while expected economic return (U_{total}) is *negative* for the first 2 years and below the level of strategic need (U_{target}) throughout the period A to C. Expected economic return (U_{total}) changes over time in accordance with the common finding that newly hired individuals do not immediately yield high steady-state performance levels (cf. Deadrick & Madigan, 1990).

Second, the firm will net a positive economic return from the selection system if the area labeled "AHG" is less than the area labeled "GDF" (we assume that all the appropriate discounting, taxes, etc. have been applied to all future economic returns). Third, and most importantly, the firm will have met its strategic goal if the area labeled "CDE" is greater than the area labeled "ABC." This example provides a straightforward extension of Schmidt et al.'s (1979) and Cronshaw and Alexander's (1985, 1991) applications of the BCG model using a comparison of areas above and below U_{total} (and not $\Delta U_{selection}$) as the basis of a decision to implement the selection system. Specifically, decision makers following the Schmidt et al. and Cronshaw and Alexander procedures would have combined future values of $\Delta U_{selection}$ (Equation 3) to arrive at an estimate of "return" which, while extremely (and accurately) large, is not a realistic or valid estimate of how well strategic need is met (i.e.,

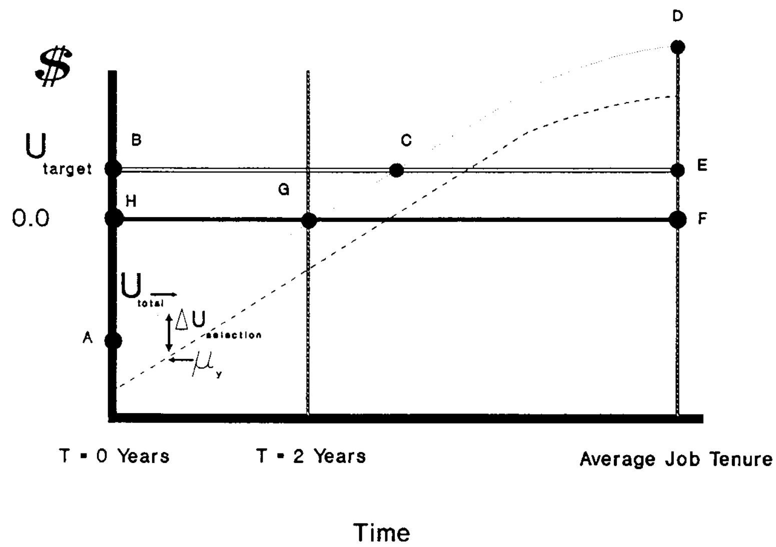
U_{total} relative to U_{target} in each future time period). The next example considers a situation where U_{target} is not constant over time.

An Undercapitalized Entrepreneur

Let us change the strategic scenario by making our firm an exploiter of new product and market opportunities. This firm also generates revenue through retail sales, though primarily through the "innovative" means of leveraging growth in market share using franchisee capital (a developing trend in the 1990's U.S. economy, Martin, 1988). We will assume that the franchisees own and operate independent restaurants. The labor market (applicants for franchises) faced by this firm consists of individuals both in and out of the restaurant industry who want to own and operate their own business. Our hypothetical firm intends to use a cognitive ability test to select franchisees. However, this industry faces a labor market characterized by applicants with a lot of desire and little capital, causing the failure of many new franchisees as initial cash flow and profits are inadequate to meet the personal needs of the franchisee during the start-up period (Evans & Jovanovic, 1989; Martin, 1988). If we assume that the average franchisee has capitalization to last one year before showing a profit, the selection instrument that predicts U_{total} to be negative until year 2 is of no value to our restaurant franchiser (see Figure 1). Since the average franchisee is predicted to go bankrupt at the end of year 1, fixed costs incurred for franchisee training and start-up costs yield a negative expected total utility to the franchiser.

Decision Implications of a Strategic Focus

Stated simply, in the economic survival (Russell & Domm, 1990) and "status quo" examples, the firm needs to ask: (a) Given all other factors besides the selection system (e.g., capitalization, availability of raw materials, etc.), what is the expected level of performance generated by a manager? (b) How much of a performance increment could be expected if a new selection system is used? and (c) Are the levels of performance expected with or without the selection system adequate to meet the firm's strategic need? After addressing these questions, an additional question found at the heart of traditional applications of BCG utility estimates must be asked: Is the incremental performance increase expected from selection instrument "A" greater than that expected from instrument "B"? The examples differ from traditional BCG applications only in that the focus has shifted from $\Delta U_{selection}$ to include the comparison of U_{total} to U_{target}.



U_{target} = the level of utility (value) dictated by the firm's strategic goals
G = "break even" point, where cost of labor = total utility derived
C = point at which total utility derived from labor = strategic target
If Area AHG < Area GDF, then positive economic return
If Area CDE > Area ABC, then strategic goal is met

Figure 1: **Utility derived from selection ($\Delta U_{selection}$) and total utility derived from human resources (U_{total}) over time: $\Delta U_{selection} = U_{total} - \mu_y$**

The restaurant example introduces an additional source of variation in that strategic need is now a moving target. If strategic objectives are to be met in the face of limited access to capital, the required "alignment" of resources (capital and human) must be altered. More human resources in the form of higher skill levels must be brought into the production process at an earlier point in time. Hence, the answer to the question regarding whether the expected performance levels are adequate to meet the firm's needs becomes complicated. Instruments with positive selection utility may not meet the strategic needs of the firm in any of the examples, though the timing of the restauranteur's strategic needs may add an additional reason for the rejection of an instrument, namely, that it does not deliver the essential levels of performance at the right time. In sum, depending on the strategic needs of the firm (e.g., survival, maintenance of status quo, or market penetration and growth), differences between the traditional use of $\Delta U_{selection}$ and our weighted combinations of U_{total} matter.

Note that our restauranteur requires performance for the first year to be at some specified level for the continued fiscal health of the franchisee. The strategic decision maker at our restaurant franchiser would be using a conjunctive decision model (Einhorn, 1970) to decide whether the selection system will generate adequate economic return during the critical initial planning period. It is a conjunctive decision process under conditions where U_{total} must be positive in a "single hurdle" decision model, that is, where U_{total} must be positive after one year. Assuming fixed costs for labor and the selection system, either a better test will need to be acquired (i.e., increase r_{xy}), the labor pool will need to be enhanced (increase μ_y, as per Boudreau & Rynes, 1985), or both. Regardless, while derivations of the basic BCG formulae are used to arrive at estimates of utility in each interval of the planning period, the firm's strategic target will dictate how these estimates should be combined in deciding whether the selection system contributions are adequate. Our "status quo" firm simply adds them up, while our restaurant franchiser must consider a nonlinear, conjunctive configuration.

Implications for Estimating How Well Strategic Need Is Met: Changing Values of r_{xy} and SD_y Over Time

Hunter, Schmidt, and Coggin (1988) recently criticized capital budgeting methods for utility estimates that appear extreme compared to competing investment decisions. Such extreme values make utility estimates of personnel interventions appear so preposterous that they are dismissed as ridiculously inaccurate by managerial decision makers (cf. Ashe, 1990). Ashe suggested that estimates of utility which dwarf the national debt (e.g., Hunter & Schmidt, 1982) preclude their credible use in management decisions *and* in arguments of legal defensibility. Unfortunately, even in applications of the BCG model that have been adjusted with capital budgeting procedures, utility estimates tend to be unrealistically large.

One reason current return on investment (ROI) estimates derived from selection utility (Cronshaw & Alexander, 1991) may seem inflated relative to estimates of ROI made for competing investments (e.g., retooling, acquisition of new retail outlets) is the assumption that "things stay the same." Boudreau (1983a) speculated whether effects of human resource interventions (e.g., r_{xy}) remain constant for each relevant performance measure over some planning period *and* whether estimates of SD_y in the applicant pool would vary as the labor market changed over time throughout the planning period. Hence, as with our demonstration of how U_{target} can change over time, U_{total} and $\Delta U_{selection}$ may vary over time as estimates of r_{xy} and SD_y change. Instead of single values of r_{xy} and SD_y derived from predictors measured at T_1 and a criterion measured at T_2, the BCG model must be modified to account for multiple criterion measures taken at T_2, T_3, $T_4 \ldots T_n$. Hunter et al. (1988) and Cronshaw and Alexander (1985, 1991) did consider utility gained at multiple points in the future; however, they assumed r_{xy} and SD_y remain constant such that expected returns in each future period are the same. Regardless, almost all derivations of selection utility over time can be modified to accommodate changes in r_{xy} and SD_y (e.g., Boudreau, 1983a, Equation 11, or any of Cronshaw & Alexander's, 1985, capital budgeting derivations).

A large literature suggests that r_{xy} is not constant over time (cf. Fleishman & Quaintance, 1984), while only one study has reported empirical evidence of instability in SD_y over time (Deadrick & Madigan, 1990). We present a brief review of the evidence suggesting longitudinal instability in r_{xy} and SD_y, and summarize how estimates of selection utility ($\Delta U_{selection}$) and total utility (U_{total}) might change.

Stability of r_{xy} and SD_y

Stability of r_{xy}. Fleishman and his colleagues have spent almost 40 years examining longitudinal relationships between predictor and criterion construct domains (Fleishman & Quaintance, 1984). Fleishman and Hempel (1954, 1955), Fleishman (1960), and Fleishman and Rich (1963) have demonstrated that: (a) Cognitive ability tends to be highly correlated with initial task performance and *decreases* in correlations with subsequent performance trials until zero or nonsignificant criterion-related validities result, and (b) motor response time and perceptual motor skill tests tend to be uncorrelated with initial performance levels and *increase* in correlation with later task performance. Though most of these findings were obtained in laboratory settings using relatively simple tasks, they underscore the need to examine criterion performance measures taken at more than one point in time.

Hulin, Henry, and Noon (1990) examined the temporal trends in predictive validity across 41 studies and 77 independent validity sequences. Predictors spanned a broad range of cognitive and psychomotor tests (e.g., from Law School Admissions Test and undergraduate GPA to rotary pursuit tasks and two-handed coordination tests). Hulin et al. found substantial evidence of decreasing trends in validity over time across all tasks and/or jobs. Their summary evidence strongly suggests that: (a)

The combinations of abilities required for job performance change over time, and/or (b) incumbents' skills and abilities change with practice.

Conclusions from these findings are far from universally accepted (see Austin, Humphreys, & Hulin, 1989; Barrett & Alexander, 1989; Barrett, Alexander, & Doverspike, 1992; Barrett, Caldwell, & Alexander, 1985; Deadrick & Madigan, 1990; Henry & Hulin, 1989; Hoffman, Jacobs, & Gerras, 1992, for contrasting views). Ackerman (1986, 1987, 1989) proposed that observed changes in relationships between cognitive and psychomotor skill tests and task performance over time can be explained by a moderator variable—the cognitive processing requirements of the task. Ackerman used the distinction between automatic and controlled cognitive processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) to explain why relationships between (a) cognitive and psychomotor predictors and (b) task performance criterion measures increase, decrease, or are sustained over time. He hypothesized that measures of cognitive skills should exhibit little or no decay in predictive validity when tasks are inconsistent in their "rules for information processing... components of processing, or... sequences of information processing components" (Ackerman, 1987, p. 4). Performance of such tasks is hypothesized to be dependent on the limitations of an individual's cognitive resources. Hence, measures of those cognitive resources should exhibit sustained predictive validity with task performance over time. This explanation is consistent with Hunter's (1986) observation that measures of general cognitive ability (*g*) tend to yield higher criterion-related validities in complex jobs (Hunter's term "job complexity" implies a continuum of jobs that vary in quantity of consistent vs. inconsistent tasks). Hunter (1986) speculated that this was due to complex jobs requiring incumbents to innovate (a complex cognitive task) when faced with "situations that will be only approximately like the situation described when the rules were given" (p. 348). It is also interesting to note that the School and College Ability Test used in the Management Progress Study at AT&T demonstrated an increase in criterion-related validity between the 8- and 20-year time lag (Bray, Campbell, & Grant, 1974; Howard & Bray, 1988), a finding that was not included in the Hulin et al. (1990) review.

In contrast, automatic processes are thought to develop over time on tasks that are easily broken into units and procedures, limiting the amount of controlled monitoring and information processing that is needed. These tasks may require extensive controlled processing initially. However, the quantity of cognitive resources required decreases over time. Consequently, measures of *g* will demonstrate criterion-related validity with early trials of task performance that slowly shrink over task repetition as automatic processes reduce the need for controlled cognitive resources. Perceptual or psychomotor measures of more specialized skills would be expected to demonstrate sustained or increasing criterion validities with performance on more consistent tasks.

Hence, there is reason to believe that, at least in the task-job combinations examined by Fleishman and Ackerman, estimates of r_{xy} are not constant over time for two common classes of predictors: tests of general intellectual ability and psychomotor skill. Evidence of substantial changes in r_{xy} is prominent with other commonly used selection systems (cf. Bray et al., 1974; Howard & Bray, 1988). The implications derived for forecasts of selection utility and total utility are straightforward but substantial. It would be predicted that criterion-related validity for cognitive ability and perceptual/psychomotor tests will be moderated by the automatic versus controlled cognitive requirements found in the criterion construct space. Other things being equal, criterion-related validities with long-term performance (and hence, utility) should be augmented most by perceptual/psychomotor predictor tests in jobs that permit automatic functioning for the majority of the applicant pool after some initial on-the-job learning period. In contrast, general cognitive ability predictor tests should yield consistently high validities with long-term performance in jobs that require controlled cognitive functioning.

Variations in SD_y over time. It is instructive to revisit the definition of SD_y before exploring the possibilities for true change in SD_y over time. SD_y is the standard deviation of performance measured in dollar terms that one would observe if the entire applicant pool were hired. Further, any estimate of SD_y will be specific to the time at which the criterion measure Y is obtained. Consequently, if an entire applicant cohort is hired and 20 criterion measures are obtained over the first 20 months of employment, 20 different estimates of r_{xy} *and* SD_y can be obtained.

While Landy, Farr, and Jacobs (1982) and others have called for research on differences in SD_y across jobs, job families, and industry, we are unaware of any studies of SD_y in which an entire applicant cohort is hired with absolutely no screening procedure of any kind. To our knowledge, only one study has reported actual change in SD_y over time in a sample that was *not* prescreened using the selection instrument of interest. Deadrick and Madigan (1990) reported increases of 7%, 18%, and 17% in SD_y estimates over a 24-week period among three samples of sewing machine operators. These differences translate into 7%, 18%, and 17% changes in $U_{selection}$ (percentage changes in U_{total} will also depend on the base-line level of performance, $N_{employed}[\mu_y - C_{labor}]$). Interestingly, Deadrick and Madigan (1990) found average hourly production rose from approximately $2.00 to $4.50 per hour over a 24-week period, consistent with our example in Figure 1.

Though changes in SD_y are not typically studied, the pattern of decreasing correlations between performance measures as time lags between measures increase (commonly called the simplex pattern) has been repeatedly observed over the last 35 years (Ackerman, 1987; Henry & Hulin, 1987). One straightforward interpretation of the simplex pattern is that variance as reflected in subjects' rank order in early job-performance measures tends not to be related to variance in later job-performance measures. While the simplex pattern of correlations suggests that individuals' relative performance levels change over time, it may be that total variation around mean performance stays the same. Hence, SD_y at time T_1 may be the same as SD_y at time T_2. However, this seems unlikely, especially in jobs where new hires have no prior experience and must learn on the job. New hires' average performance level ($\overline{Y}_{selectee}$) would be expected to be low *initially*, as would their SD_y (i.e., a tight distribution of performance measures at the low end of the scale). It would also be expected that, over time, average performance level would increase. SD_y might also increase as individual differences emerge. These individual differences may be in *g*, or in levels of motivation and/or opportunity. Alternatively, in jobs that are machine-paced, SD_y might be initially large, then shrink over time as members of the cohort group reach some machine-dictated asymptotic performance level. Regardless of whether the focus is on new hires with no prior experience or trained professionals expected to step in and perform immediately, it seems unlikely that variation in contributions to the firm will remain constant over the employment tenure of a cohort of newly selected employees, though the exact nature of such variations remains to be understood.

Summary and Example

Importantly, selection utility *and* total utility derived from the selection tests as calculated when r_{xy} and SD_y vary over time will be substantially different from traditional applications of the BCG model. The impact of the two utility parameters r_{xy} and SD_y on selection utility and total utility can only be derived if longitudinal changes in r_{xy} and SD_y are known. If these changes *were* known, they could be used to evaluate the "risk" associated with each selection battery/job combination—those combinations with greater temporal variation in r_{xy} and SD_y are a "riskier" investment for the firm (financial theory dictates that riskier investments be paired with higher discount rates as described in Boudreau's, 1983a, and Cronshaw & Alexander's, 1985, 1991, capital budgeting adjustments to BCG models).

Estimates of selection and total utility become further complicated when batteries of tests are used, where all predictors are correlated with one another *and* exhibit different rates of change in r_{xy} as the predictor-criterion time lag increases (sample derivations are available from the first author). As an example, suppose X_1 is the U.S. Office of Personnel Managements' General Aptitude Test Battery (GATB) that predicts initial on-the-job performance (Y_3) with a typical validity coefficient of .40 (i.e., $r_{1,3} = .40$). Also assume that X_1 predicts later on-the-job performance (Y_4) with validity .30 ($r_{1,4} = .30$). Now, suppose that a psychomotor test (X_2) is correlated with initial performance at $r_{2,3} = .20$. However, Ackerman's findings suggest that this validity may increase substantially over time for "automatic" jobs. So, for example, we let $r_{2,4} = .60$. Finally, we let the correlation between the two predictors be .30 ($r_{1,2} = .30$). Assuming $SD_3 = SD_4$ and all other aspects of the utility model are fixed (e.g., selection ratio, costs of selection, etc.), the utility of GATB and psychomotor tests on initial performance is driven by the product $r_{xy} \cdot SD_y$. The value of $r_{xy} \cdot SD_3$ is:

$$\frac{(r_{1,3} + r_{2,3})SD_3}{\sqrt{2 + 2r_{1,2}}} = \frac{(.4 + .2)SD_3}{\sqrt{2.6}} = .37SD_3,$$

while the product generated for later performance (Y_4) is:

$$\frac{(r_{1,4} + r_{2,4})SD_4}{\sqrt{2 + 2r_{1,2}}} = \frac{(.3 + .6)SD_4}{\sqrt{2.6}} = .56SD_4,$$

(Derivations of these formulae are available from the first author). Because $SD_3 = SD_4$, the utility obtained (before considering costs or depreciation) from later on-the-job performance is $(.56 \div .37)100 = 150\%$ greater than the utility obtained from early on-the-job performance. Clearly, differences in criterion-related validities over time within a multiple-test selection battery combined with changes in SD_y can meaningfully impact forecasted selection utility.

Discussion

Throughout this paper we have adopted the perspective of someone who is required to gather relevant information before embarking the firm on some course of action (i.e., strategic decision makers and human resource managers). A criterion measure is needed to help the strategic decision maker determine whether alternate selection systems contribute more or less to the firm's strategic objectives. We argued that the BCG model of selection utility as it is commonly used is deficient for this purpose. First, a metric of utility relative to random selection, pegged to a single point in time at which a lone performance measure was taken, does not tell strategic managers what they need to know. Specifically, strategic decision makers need to know whether the expected total utility forecast by a given "alignment" of resources, in this case the alignment of human resources generated by a selection instrument(s), will meet the firm's strategic needs. We have shown that a strategic perspective dictates consideration of sound estimates of U_{total}.

Second, recent findings regarding the stability of r_{xy} (Hulin et al., 1990) and SD_y (Deadrick & Madigan, 1990) strongly indicate that changes in r_{xy} and SD_y are likely to occur over time. It is clear how temporal changes in r_{xy} and SD_y impact the validity of $U_{selection}$ and U_{total} estimates. Strategic decision makers require realistic, valid forecasts of total expected utility that account for known changes in relevant parameters. While previous investigators have permitted consideration of changes in tax rates, discount rate, and testing costs, the impact of true temporal changes in two fundamental parameters (r_{xy} and SD_y) has not been explored. Variations in r_{xy} and SD_y will also be important to the human resource decision makers who use estimates of $\Delta U_{selection}$ to decide among alternate expenditures from limited personnel budgets. Interestingly, by augmenting the decision context to include strategic issues, we arrive at a somewhat different emphasis than Boudreau (1991). He argued that SD_y estimation research "can advance measurement theory" and that the "value of [SD_y] research rests *not* on its ability to better describe, predict, explain, or enhance decisions" (emphasis added, p. 660). Rather, we show that valid point estimates of SD_y, r_{xy}, and their changes over time become critical when strategic issues are considered. Any representation of selection utility that does not accommodate changes in firms' strategic needs *and* changes in the capacity of a selection system to meet those needs (i.e., changes in r_{xy} and SD_y that impact U_{total} and $U_{selection}$) will be deficient.

Many directions for future research are implied by our analysis. Examinations of the dynamic nature of r_{xy} initiated by Ackerman (1987); Alexander, Barrett, and Doverspike (1991); and Deadrick and Madigan (1990) need to be extended. Unfortunately, similar efforts exploring the nature of changes in SD_y across time seem to be critical, yet almost nonexistent. Further, estimates of SD_y derived from nonmarket sources (e.g., subject matter expert judgments) are prone to distortion (Bobko, Karren, & Kerkar, 1987). Shetzer and Bobko (1987) and Bobko, Shetzer, and Russell (1991) demonstrated that the stake-holder group used to generate SD_y estimates (i.e., customers vs. supervisors) will greatly impact both the estimates and the types of error in those estimates. It is routinely assumed by strategic decision makers that stake-holders from the investor community, management, incumbents, customers, and others attach different value to job performance and hence yield different estimates of SD_y. Currently, the most consistent SD_y estimates seem to be derived from labor market wage and salary surveys. Unfortunately, this stability may be deceiving. Specifically, estimates of SD_y derived from labor market wage and salary surveys can be prone to nonrandom contamination (see England & McLaughlin, 1979, and England, 1982, for a description of how barriers to entry distort market forces, specifically labor supply, behind "equilibrium" wages). Exploration of alternate, nondollar means of calibrating SD_y have taken place in public sector and military settings (Eaton, Wing, & Lau, 1985; Eaton, Wing, & Mitchell, 1985). It remains to be seen whether these methods are less prone to systematic and random measurement error.

Interestingly, we are aware of no research examining the various influences on μ_y other than Boudreau and Rynes' (1985) examination of differences in initial values of μ_y achieved through alternate recruiting procedures for sampling applicants from the labor pool. As noted above, this estimate has a major impact on the critical difference between notions of selection utility and total utility, yet no literature examines the relative merits of (a) implementing a selection system for an existing job and its applicant pool versus (b) redesigning the job to obtain higher expected economic returns from the combination of a different selection system and different applicant pool. On a broader scale, estimates of μ_y for regional or national labor markets should factor into decisions concerning where facilities should be located, under the assumption that in a global economy all aspects of production (capital, technology, raw materials, etc.) are mobile *except* labor (in the short term). Finally, as is implicit in Figure 1 and documented by Deadrick and Madigan (1990), μ_y will change for a cohort group over time through maturation and learning acquired on-the-job. Knowledge of the initial values of μ_y, r_{xy}, and SD_y that characterize each labor market *and how they change over time* will be critical in determining the expected economic return from each location.

In sum, decisions based on the financial impact of selection systems, indeed, all human resource management systems, take place in a richer decision context than heretofore described. We have tried to refocus attention on the underlying economic criterion construct domain by partitioning the utility construct into notions of selection utility and total utility. The impact of total human resource utility on accomplishment of strategic needs was delineated. We further questioned the viability of implicitly assuming that key parameters (r_{xy} and SD_y) remain constant over time. Hence, for selection research to have its proper impact on managers' decisions, we must (a) expand the focus of our efforts to include estimates of absolute economic return that can be used by strategic managers, and (b) reassess our assumptions concerning the temporal stability of parameter estimates.

REFERENCES

Ackerman PL. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence*, *10*, 101–139.

Ackerman PL. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, *102*, 3–27.

Ackerman PL. (1989). Within-task intercorrelations of skilled performance: Implications for predicting individual differences? (A comment on Henry & Hulin, 1987). *Journal of Applied Psychology*, *74*, 360–364.

Alexander RA, Barrett GV, Doverspike D. (1991, April). Appropriate (and inappropriate) data and analyses relevant to the dynamic criterion question. In Hanges PJ (Chair), *Dynamic criteria: Are we shooting at a moving target?* Symposium conducted at the Sixth Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Ansoff HI. (1988). *The new corporate strategy*. New York: Wiley.

Ashe RL Jr. (1990, April). *The legal defensibility of assessment centers and in-basket exercises*. Presented at the Fifth Annual Conference of the Society for Industrial and Organizational Psychology, Miami Beach, FL.

Austin JT, Humphreys LG, Hulin CL. (1989). Another view of dynamic criteria: A critical reanalysis of Barrett, Caldwell, and Alexander. PERSONNEL PSYCHOLOGY, *42*, 583–596.

Barrett GV, Alexander RA. (1989). Rejoinder to Austin, Humphreys, and Hulin: Critical reanalysis of Barrett, Caldwell, and Alexander. PERSONNEL PSYCHOLOGY, *42*, 597–612.

Barrett GV, Alexander RA, Doverspike D. (1992). The implications for personnel selection of apparent declines in predictive validities over time: A critique of Hulin, Henry, and Noon. PERSONNEL PSYCHOLOGY, *45*, 601–618.

Barrett GV, Caldwell MS, Alexander RA. (1985). The concept of dynamic criteria: A critical reanalysis. PERSONNEL PSYCHOLOGY, *38*, 41–56.

Bobko P, Karren R, Kerkar SP. (1987). Systematic research needs for understanding supervisory-based estimates of SD_y in utility analysis. *Organizational Behavior and Human Decision Processes*, *40*, 69–95.

Bobko P, Shetzer L, Russell CJ. (1991). On the robustness of judgments of SD_y in utility analysis: Effects of frame and presentation order. *Journal of Occupational Psychology*, *64*, 179–188.

Boudreau JW. (1983a). Economic considerations in estimating the utility of human resource productivity improvement programs. PERSONNEL PSYCHOLOGY, *36*, 551–557.

Boudreau JW. (1983b). Effects of employee flows on utility analysis of human resource productivity improvement programs. *Journal of Applied Psychology*, *68*, 396–407.

Boudreau JW. (1991). Utility analysis for decisions in human resource management. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 621–745). Palo Alto, CA: Consulting Psychologists Press.

Boudreau JW, Berger CJ. (1985). Decision-theoretic utility analysis applied to employee separations and acquisitions [Monograph]. *Journal of Applied Psychology*, *70*, 581–612.

Boudreau JW, Rynes SL. (1985). Role of recruitment in staffing utility analysis. *Journal of Applied Psychology, 70*, 354–366.
Bray DW, Campbell RJ, Grant DL. (1974). *Formative years in business: A long-term AT&T study of managerial lives*. Malabar, FL: Krieger.
Brogden HE. (1946). On the interpretation of the correlation coefficient as a measure of predictive efficiency. *Journal of Educational Psychology, 37*, 65–76.
Brogden HE. (1949). When testing pays off. PERSONNEL PSYCHOLOGY, *2*, 171–183.
Cascio WF. (1991). *Costing human resources: The financial impact of behavior in organizations* (3rd ed.). Boston: Kent.
Cronbach LJ, Gleser G. (1965). *Psychological tests and personnel decisions*. Urbana, IL: University of Illinois Press.
Cronshaw SF, Alexander RA. (1985). One answer to the demand for accountability: Selection utility as an investment decision. *Organizational Behavior and Human Decision Processes, 35*, 102–118.
Cronshaw SF, Alexander RA. (1991). Why capital budgeting techniques are suited for assessing the utility of personnel programs: A reply to Hunter, Schmidt, and Coggin (1988). *Journal of Applied Psychology, 76*, 454–457.
Cronshaw SF, Alexander RA, Wiesner WH, Barrick MR. (1987). Incorporating risk into selection utility: Two models for sensitivity analysis and risk simulation. *Organizational Behavior and Human Decision Processes, 40*, 270–286.
Deadrick DL, Madigan RM. (1990). Dynamic criteria revisited: A longitudinal study of performance stability and predictive validity. PERSONNEL PSYCHOLOGY, *43*, 717–744.
Devanna MA, Frombrun C, Tichy N. (1981). Human resources management: A strategic perspective. *Organizational Dynamics, 9*(3), 51–67.
Eaton NK, Wing H, Lau A. (1985, November). *Utility estimation in five enlisted occupations*. Paper presented at the 1985 Military Testing Association Meeting, San Diego, CA.
Eaton NK, Wing H, Mitchell KJ. (1985). Alternative methods of estimating the dollar value of performance. PERSONNEL PSYCHOLOGY, *38*, 27–40.
Einhorn HJ. (1970). Use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin, 73*, 221–230.
England P. (1982). Do men's jobs require more skills than women's? *ILR Report, 19*(2), 20–23.
England P, McLaughlin SD. (1979). Sex segregation of jobs and male-female income differentials. In Alvarez R (Ed.), *Discrimination in organizations*. San Francisco, CA: Jossey-Bass.
Evans DS, Jovanovic B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy, 97*, 808–827.
Fleishman EA. (1960). Abilities at different stages of practice in rotary pursuit performance. *Journal of Experimental Psychology, 60*, 162–172.
Fleishman EA, Hempel WE Jr. (1954). Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika, 19*, 239–252.
Fleishman EA, Hempel WE Jr. (1955). The relation between abilities and improvement with practice in a visual discrimination reaction task. *Journal of Experimental Psychology, 49*, 301–316.
Fleishman EA, Quaintance MK. (1984). *Taxonomies of human performance: The description of human tasks*. Orlando, FL: Academic Press.
Fleishman EA, Rich S. (1963). Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *Journal of Experimental Psychology, 66*, 6–11.
Hamel G, Prahalad CK. (1989). Strategic intent. *Harvard Business Review, 67*, 63–76.
Henry RA, Hulin CL. (1987). Stability of skilled performance across time: Some generalizations and limitations on utilities. *Journal of Applied Psychology, 72*, 457–462.

Henry RA, Hulin CL. (1989). Changing validities: Ability-performance relations and utilities. *Journal of Applied Psychology*, *74*, 365–367.

Hoffman DA, Jacobs R, Gerras S. (1992). Mapping individual performance over time. *Journal of Applied Psychology*, *77*, 185–195.

Howard A, Bray DW. (1988). *Managerial lives in transition: Advancing age and changing times*. New York, NY: Guilford Press.

Hulin CL, Henry RA, Noon SL. (1990). Adding a dimension: Time as a factor in generalizability of predictive relationships. *Psychological Bulletin*, *107*, 328–340.

Hunter JE. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*, 340–362.

Hunter JE, Schmidt FL. (1982). Fitting people to jobs: The impact of personnel selection on national productivity. In Dunnette MD, Fleishman EA (Eds.), *Human performance and productivity*, (Vol. 1, pp. 233–284). Hillsdale, NJ: Erlbaum.

Hunter JE, Schmidt FL, Coggin TD. (1988). Problems and pitfalls in using capital budgeting and financial accounting techniques in assessing the utility of personnel programs. *Journal of Applied Psychology*, *73*, 522–528.

Landy FJ, Farr JL, Jacobs RR. (1982). Utility concepts in performance measurement. *Organizational Behavior and Human Performance*, *30*, 15–40.

Martin RE. (1988). Franchising and risk management. *American Economic Review*, *78*, 954–968.

Murphy KR. (1986). When your top choice turns you down: Effect of rejected offers on the utility of selection tests. *Psychological Bulletin*, *99*, 133–138.

Rich JR, Boudreau JW. (1987). The effects of variability and risk in selection utility analysis: An empirical comparison. PERSONNEL PSYCHOLOGY, *40*, 55–84.

Russell CJ, Domm DR. (1990, August). *On the validity of role congruency-based assessment center procedures for predicting performance appraisal ratings, sales revenue, and profit*. Presented at the Fifty-Fourth Annual Meeting of the Academy of Management, San Francisco.

Schmidt FL, Hunter JE, McKenzie RC, Muldrow TW. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, *64*, 609–626.

Schneider W, Shiffrin RM. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Schuler RS, Jackson SE. (1987). Linking competitive strategies with human resource management practices. *Academy of Management Executive*, *1*, 207–219.

Shetzer L, Bobko P. (1987, August). *The effects of frame and anchoring on estimates of overall worth in utility analysis*. Paper presented at the Forty-Seventh Annual Meeting of the Academy of Management, New Orleans, LA.

Shiffrin RM, Schneider W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Smith PC. (1976). Behaviors, results, and organizational effectiveness: The problem of criteria. In Dunnette MD (Ed.) *Handbook of industrial and organizational psychology* (pp. 745–766). Chicago: Rand McNally.

Snow CC, Snell SA. (1993). Staffing as strategy. In Schmitt N, Borman WC (Eds.), *Personnel selection in organizations*. San Francisco: Jossey-Bass.

Tsui AS, Gomez-Mejia LR. (1988). Evaluating human resource effectiveness. In Dyer L, Holder GW (Eds.), *Human resource management: Evolving roles and responsibilities* (pp. 1-187-1-227). Washington, DC: Bureau of National Affairs.