
Utility Analysis: Current Trends and Future Directions

Elizabeth F. Cabrera and Nambury S. Raju*

Utility analysis procedures offer organizational decision makers useful information regarding the relative values of different interventions. Years of research have resulted in a number of practically viable utility models and extensions. There is a continued need for research to examine the accuracy of utility estimates and to further compare the different models. A more recent research concern is that of low levels of acceptance of utility analysis results by practitioners. Many researchers are turning their attention to ways in which this acceptance may be increased. This article reviews different utility models as well as a number of important extensions. It then discusses current utility analysis issues, such as the aforementioned acceptance problem and the introduction of a multi attribute utility model. The article concludes with suggestions for future utility analysis research.

Introduction

Utility analysis (UA) refers to a set of methods and procedures that help organizational decision makers choose among alternatives by providing information regarding the relative values of different interventions. UA was introduced more than half a century ago (Brogden 1949). However, it was used very little during the next thirty years due to problems associated with the estimation of a crucial variable in the model, namely the standard deviation of job performance in monetary units (SD_Y). Interest in UA was renewed with the development of a method for estimating SD_Y (Schmidt, Hunter, McKenzie and Muldrow 1979).

Since then there has been a steady interest in UA research. Other methods have been proposed for estimating SD_Y (Cascio and Ramos 1986; Hunter and Schmidt 1982), Brogden's original UA model has been modified and extended (Cronbach and Gleser 1965; Boudreau, 1983a, 1983b; Boudreau and Berger, 1985a, 1985b; De Corte 1994, 1996, 1998) and new UA models have been developed (Raju, Burke and Normand 1990; Raju, Cabrera and Lezotte 1996).

Despite these advances in UA methods, organizational decision makers have resisted the adoption of UA procedures in practice. In response to this problem, researchers have turned their attention to issues of acceptance of UA recommendations by managers (Carson, Becker and Henderson 1998; Cronshaw 1997; Hazer and Highhouse 1997; Latham and Whyte 1994; Roth, Segars and Wright, 2000; Whyte and Latham 1997) and to multi attribute UA models (Roth and Bobko 1997).

Following a description of Brogden's original utility model and a careful examination of the key variables, each of the formerly mentioned contributions to UA research will be discussed. The article then concludes with some recommendations for future research directions. Given the nature of this special issue, the article will focus on the utility of selection programs, although the models and procedures reviewed here are equally valid for other organizational interventions such as training.

Brogden-Cronbach-Gleser Model

Most of the current work on UA is based on the classical utility model developed by Brogden (1949). While other models existed for estimating the utility of selection procedures (Sands 1973; Taylor and Russell 1939), none allowed for the expression of utility in monetary terms. Brogden proposed a formula to transform the validity coefficient of a selection program into an estimate of monetary value. His formula was based on the assumption of a linear relationship between a predictor score and the monetary value of performance.

Cronbach and Gleser (1965) expanded Brogden's model to include the cost of testing applicants. The resulting Brogden Cronbach Gleser (BCG) model, which shows the incremental utility or productivity gain of a predictor based selection procedure over random selection when N_s applicants are hired, can be expressed as:

$$\Delta U = N_s SD_Y (r_{XY})(\mu_{X_s}) \quad NC \quad (1)$$

* Address for correspondence:
Departamento de Economía de
la Empresa, Universidad Carlos
III de Madrid, c/ Madrid 126,
28903 Getafe (Madrid), Spain.
E-mail: beth@emp.uc3m.es

where SD_Y is the standard deviation of job performance in monetary units (Y), r_{XY} is the correlation between the predictor (X) and $Y\mu_{X_s}$, is the mean predictor score for the selectees, N is the total number of applicants and C represents the average cost per applicant of administering the selection procedure.

In addition to the assumption of a linear relationship between X and Y , this equation assumes (mostly for convenience) that the predictor score X is standardized, i.e. has a mean of zero and a standard deviation of one. Two additional assumptions allow for changes in the equation that make it easier to apply in practice. First, it is assumed that the correlation between X and Y is equal to the correlation between X and R , a rating of job performance. Thus, the known value of the correlation between the predictor score and a rating of job performance can be substituted for the unknown correlation between the predictor and the monetary value of performance. Second, the assumption that the predictor score X is normally distributed and that selectees are hired top down makes the direct calculation of μ_{X_s} unnecessary. Given these assumptions, the gain in utility may be expressed as:

$$\Delta U = N_s SD_Y (r_{XR}) \frac{\phi}{p} NC \quad (2)$$

where r_{XR} is the correlation between the predictor and a rating of performance, p is the proportion of applicants selected (top down) or the selection ratio and ϕ is the normal ordinate associated with p .

It should be noted that these assumptions may not always hold. Murphy (1986) presented empirical evidence to show that the top applicants do not always accept employment offers. When this occurs, the true selection ratio will be different given that the selectees are not those with the highest predictor scores. On the other hand, one of the more critical assumptions of a linear relationship between X and Y has been supported empirically (Hunter and Schmidt 1982).

As mentioned previously, UA was developed as a decision aid to help professionals choose among alternatives. Therefore, it is most often used in situations where a decision must be made between two different selection procedures. The gain in utility of a new selection program over that of an old program rather than over random selection may be written as follows:

$$\Delta U = N_s SD_Y (r_2 - r_1) \frac{\phi}{p} N(C_2 - C_1) \quad (3)$$

where r_2 and r_1 represent the correlations between X and Y for the new (2) and old (1)

programs and, likewise, C_2 and C_1 represent the average cost of administering the selection procedure per applicant for the new and old programs.

Key Variables

According to the BCG utility model, as Equation 1 shows, the utility of any selection program depends on five key variables (Cabrera 1998). These five variables are briefly described below.

Number of applicants accepted. Other things being equal, the more applicants selected using a particular selection procedure, the greater the total utility. How many are hired at a given time depends upon the needs of an organization. For example, some organizations may be growing faster than others and are, thus, more likely to hire many new employees (and more often) than slower growing organizations. Or some areas within an organization may develop greater staffing needs due to environmental changes or strategic decisions. Also, some jobs may have higher turnover rates than others, thus requiring that new employees be hired more often. In general, the utility of a selection procedure will increase as staffing needs increase.

Standard deviation of performance in monetary units (SD_Y). The difference in standard deviations among different jobs depends on two major factors. The first factor is the nature of the job. Jobs differ in the degree to which employee ability or motivation affects performance. That is, some jobs are more sensitive to the efforts of employees than others. This will depend to a great extent on the amount of autonomy employees have regarding how to perform the work. Jobs characterized by strict routines and procedures limit employee initiative and, therefore, reduce the possibility of encountering variance due to employee differences. For example, assembly line workers are limited by the assembly process so that two workers with different abilities or motivation will still have more or less the same performance. On the other hand, the project leader of an advertising team has much more freedom to decide how to carry out the job, in which case there could be great discrepancies in the performance of two team leaders due to different levels of ability or motivation.

Hunter, Schmidt and Judiesch (1990) conducted a study in which they showed that SD_Y increases as a function of the level of job complexity. As job complexity increases it becomes more and more difficult to define procedures by which the job should be performed. Thus, it is likely that workers in these jobs have more autonomy and are less

constrained by rules and regulations. So, complex jobs should be more sensitive to employee differences and, therefore, as Hunter *et al.* (1990) found, should have greater SD_Y s.

Another factor that influences SD_Y is the relative value of a particular job to the organization. There are some jobs in which outstanding performance contributes greatly to the financial success of a firm, whereas there are other jobs in which the level of performance of the employee may have very little impact on organizational results. This depends on the relative position of the job in the organization's value chain. Some jobs are vital for realizing the strategic goals of the firm, while others are clearly less important. Although performance variance among employees in less important jobs may exist, this variance does not translate into a true difference in monetary terms for the organization. So, other variables being equal, jobs that have a larger SD_Y , either because they are more complex and permit greater employee discretion or because they have a more direct impact on organizational results, will benefit more from the use of valid selection procedures than jobs with a lower SD_Y .

Validity coefficient. Selection procedures with higher validity will have greater utility than procedures with lower validity, other things being equal. Therefore, a new selection procedure, before it is seriously considered for adoption, must have higher validity than the current selection procedure.

Selectees' mean predictor score. A higher mean predictor score will lead to a higher utility. If all of the applicants score poorly on a test intended to predict future performance, even selecting the 'best' will not contribute greatly to the financial success of the organization. Likewise, a smaller selection ratio will lead to higher utility. Given the assumption of top down selection, selecting a smaller number of the top applicants (smaller selection ratio or smaller p) will result in a higher mean test score than selecting a larger number of applicants. It should be noted, however, that in most practical situations the number of new hires will depend upon the needs of an organization. When the number of new hires is fixed, the selection ratio can only be reduced by increasing the number of applicants, which, in turn, could increase the overall cost of a new selection program by increasing recruitment costs.

Cost of the predictor and its administration. Utility will be higher, all other factors being equal, when the cost is lower. For example, the cost of a paper pencil test is cheaper than an assessment center.

Explicitly dividing the influence of utility into these five factors can be very useful for decision makers. While most understand that utility is affected by the validity of a selection procedure as well as its cost, some may not be aware that utility also depends on the job for which employees are being selected and the number and quality of the applicants. The fact that utility is influenced by the job in question has clear implications for decisions regarding selection procedures. The use of valid selection procedures for jobs that have a large SD_Y is essential. The gain in utility from replacing an old selection procedure with a new one of higher validity will be greatest for jobs with higher SD_Y .

On the other hand, gains in total utility will also be higher when the selection procedure is used to select a large number of applicants. It is often the case that the jobs requiring the greatest number of employees are those that happen to have a lower SD_Y . Thus, it is just as important to use valid selection procedures for these jobs given that the impact in total utility will be large because so many selectees are affected.

Finally, recruiting efforts are important to ensure a qualified group of applicants who will contribute to a high mean predictor score. Also, selection ratios should ideally be low, again so that the mean test score is as high as possible. As previously noted, this would require more applicants, thus increasing the cost of recruitment.

Table 1 shows the effect of differences in SD_Y , increases in the validity coefficient and the selection ratio on average utility gain per selectee, assuming cost is held constant. One can see that a new selection procedure which provides a smaller gain in validity (.25) may contribute to a higher utility gain (7,243 versus 3,139) than a procedure with a higher validity increase (.65) due to the fact that the SD_Y is greater (30,000 versus 5,000). Likewise, the

Table 1: Average gain in utility per selectee as a function of SD_Y , r_{XY} and p .

SD_Y	r_{XY}	p	$\Delta\bar{U}$
30000	.25	.40	7243
30000	.25	.70	3725
30000	.45	.40	13038
30000	.45	.70	6706
30000	.65	.40	18832
30000	.65	.70	9686
5000	.25	.40	1207
5000	.25	.70	621
5000	.45	.40	2173
5000	.45	.70	1118
5000	.65	.40	3139
5000	.65	.70	1614

average gain in utility may be higher (1,614 versus 1,207) despite a higher selection ratio (.70 versus .40), due to a much greater gain in validity (.65 versus .25).

Estimation of SD_Y

Three of the four variables called for by the BCG model are typically readily available to organizations. The only variable that is difficult for decision makers to obtain is SD_Y . Problems in estimating this variable prevented the BCG model from becoming a practically viable tool for over 25 years. Finally, Schmidt *et al.* (1979) suggested a technique for estimating the troublesome variable. Assuming that the monetary value of performance is normally distributed, they pointed out that the difference between the monetary value of an employee who performs at the 85th percentile and an employee who performs at the 50th percentile was equal to SD_Y . Likewise, the difference between an employee who performs at the 50th percentile and one who performs at the 15th percentile was also equal to the standard deviation. Therefore, Schmidt *et al.* proposed that supervisors estimate the monetary value of performance for employees at the 15th, 50th and 85th percentiles. The average of the differences between the 15th and 50th percentile estimates and the 50th and 85th percentile estimates should provide a reasonably accurate estimate of SD_Y .

Following is an example of the instructions for estimating the value of performance for a computer programmer at the 85th percentile that Schmidt *et al.* (1979) gave to supervisors in their study:

We would now like for you to consider the 'superior' programmer. Let us define a superior performer as a programmer who is at the 85th percentile. That is, his or her performance is better than that of 85% of his or her fellow programmers, and only 15% turn in better performances. Consider the quality and quantity of the output typical of the superior programmer. Then estimate the value of these products and services. In placing an overall dollar value on this output, it may again help to consider what the cost would be of having an outside firm provide these products and services. Based on my experience, I estimate the value to my agency of a superior computer programmer to be _____ dollars per year.

Similar instructions were given for the estimations of the value of performance at the 50th and 15th percentiles.

Cascio and Ramos (1986) offered an alternative method for estimating the elusive

standard deviation. Their approach, known as the Cascio Ramos estimate of performance in dollars or CREPID, requires both a job analysis and performance evaluation. The principal tasks of the job are identified and weighted according to importance. The weights are then multiplied by average salary to determine the monetary value of each principle task. These values are then multiplied by the performance rating that an employee receives for the corresponding task. The resulting figures are added to obtain the economic value of performance for an employee. The mean and standard deviation of the monetary value of performance of all employees may then be calculated (Cascio 1991).

Hunter and Schmidt (1982) also proposed two quick and easy procedures for estimating SD_Y . Based on a review of numerous studies reporting employee output data, they concluded that most estimates of SD_Y varied between 40% and 70% of mean salary. Thus, they recommend 40% and 70% of mean salary as conservative (low) and liberal (high) estimates, respectively, of SD_Y .

Several empirical studies have assessed the validity of alternative procedures for estimating SD_Y (Bobko, Karren and Kerkar 1987; Bobko, Karren and Parkington 1983; Raju, Cabrera and Lezotte 1996; Reilly and Smither 1985; Weekly, Frank, O'Connor and Peters 1985). A typical finding has been a large variation among the percentile estimates of supervisors. That is, supervisors often give very different values for the same percentile. This is not only a problem for the validity of utility estimates but may also be one of the factors which negatively affects decision makers' perception of the value of utility estimates.

Burke and Frederick (1986) proposed a modification of the Schmidt *et al.* procedure that helps to reduce the variation among percentile estimates. Following this procedure, supervisors estimate only the 50th percentile. The average of these estimates is then provided to the supervisors who go on to estimate the remaining two percentiles.

Another common finding from studies of the estimation of SD_Y is that the 40% method and CREPID tend to produce similar estimates of SD_Y , while the Schmidt *et al.* procedure tends to produce much higher estimates. This is not surprising given that both the 40% method and CREPID are based on average salary. Finally, in a study of managers' reactions to UA, Hazer and Highhouse (1997) found that in comparing the three popular methods for estimating SD_Y , managers perceived UA as being more credible and were most likely to use it when SD_Y was estimated using the 40% method.

BCG Extensions and Other Utility Models

Several extensions and variations of the BCG model have been proposed. These modifications were designed to incorporate typical business accounting practices as well as to address some of the assumptions underlying the BCG model. These are briefly described in this section.

BCG Extensions

In studying the BCG model of utility, Boudreau (1983a) concluded that key economic variables had been excluded. Specifically, the utility equation did not take into account the effects of variable costs, taxes and discounting. Not considering these economic factors, according to Boudreau, causes utility estimates to be upwardly biased. He points out that variable costs rise (fall) as productivity rises (falls). Therefore, a portion (V) of the gain in profits must be used to cover increased costs such as: variable raw materials costs, variable production overhead, benefits, and incentive or commission based pay. Likewise, a reduction in profits will be somewhat offset by a reduction in variable costs. Furthermore, a portion (TAX) of organizational profits must be used to pay taxes for those organizations facing tax liabilities. Finally, the opportunity costs of returns forgone when costs and benefits accrue over time should be accounted for by discounting the value of future costs and benefits. An extended version of the BCG equation (Equation 2) that reflects these economic considerations can be expressed as:

$$\Delta U = N_s \left[\sum_{t=1}^T \frac{1}{(1+i)^t} SD_{Y_t}(1+V_t)(1-TAX_t) \right] \left[r_{XY} \frac{\phi}{p} \right] NC(1-TAX) \quad (4)$$

where i is the discount rate, V_t is the proportion of profits represented by variable costs for time period t , TAX_t is the organization's applicable tax rate for time period t , and t is the time period in which a productivity increase occurs.

Assuming that TAX , V , SD_Y and r_{XY} remain constant over time, the gain in utility of one selection procedure over another (i.e., an extension of Equation 3) may be written as:

$$\Delta U = N_s SD_Y (1+V) (1-TAX) (r_2 - r_1) \frac{\phi}{p}$$

$$\left[\sum_{t=1}^T \frac{1}{(1+i)^t} \right] N(C_2 - C_1) (1-TAX) \quad (5)$$

In addition to suggesting the inclusion of key

economic variables in models of utility, Boudreau (1983b) also encouraged consideration of the impact of employee flow into and out of the workforce on utility estimates. Utility models had typically assumed that selection programs only affected one group of selectees. In reality, these programs are usually used numerous times, in some cases continuously, and thus affect many groups, with the impact of the programs lasting for a number of years. Employee flow is an important consideration because utility is affected by changes in the number of employees in the workforce who were chosen by a particular selection program. Boudrea (1983b) demonstrated that the use of modified utility equations which include both the effects of the three key economic variables as well as employee flow may result in greater utility estimates than previously reported.

Boudreau and Berger (1985) extended the idea of employee flow even further by considering different types of employee movement in and out of an organization. Specifically, they address the following scenarios: (1) repeated hiring without separations; (2) repeated hiring with separations not being replaced over time; and (3) repeated hiring with separations replaced over time. The last scenario is the most general one and also the most realistic from a practitioner's point of view. It allows one to see the explicit link between utility based on the BCG model and utility associated with the separation of employees for unsatisfactory performance on the job during the probationary period and replacement with new hires.

Following up on the work of Boudreau and Berger, De Corte examined, in several recent articles (1994, 1996, 1998), the question of total utility of a selection program with a probationary period. De Corte's research deals with the problem of choosing among alternative predictors for selecting a single group of new employees with the stipulation that some of these employees may be terminated later for unsatisfactory job performance. He first addresses the problem of a single cohort group (De Corte 1994) and later expands it to include multiple cohort groups as well recruitment costs (De Corte 1996). The incorporation of recruitment costs extends previous work done by Martin and Raju (1992) and Law and Myers (1993). Finally, he expands the work of Cronbach and Gleser (1965) on two stage selection to include multiple stage selection (De Corte 1998).

Other Utility Models

We now turn to two utility models developed in response to concerns about some of the assumptions underlying the BCG model.

Raju, Burke and Normand model. Raju, Burke and Normand (1990) proposed a new utility analysis model (RBN) that does not require the problematic estimation of SD_Y . Instead, the model introduces two new variables: average total compensation, A , which includes the salary, benefits, bonuses, and direct overhead allocated to an average employee and the standard deviation of a rating of job performance, SD_R . The gain in utility of a new selection procedure over an old procedure according to the RBN model may be expressed as:

$$\Delta U = N_s A (SD_R) (r_{XR}) \frac{\phi}{p} \quad NC \quad (6)$$

Although the RBN model eliminates the estimation of SD_Y , it does often require that the original performance rating scale be transformed into a new performance scale. This is necessary because the metric of any performance rating is arbitrary and may, therefore, unduly influence overall utility. Three different procedures for transforming the original performance scale were developed by Burke, Raju, Edwards and Day (1993).

Lezotte, Raju, Burke and Normand (1996) conducted a study to compare the RBN model with the BCG model. The results showed that the utility estimate derived using the RBN model was closer to a utility estimate based on an empirically derived SD_Y than was the estimate derived using a modified CREPID procedure, which was lower, and the estimate based on Schmidt *et al.*'s procedure, which was higher. In another study by Raju *et al.* (1996), utility estimates derived using the RBN model were higher than estimates based on the CREPID and the Schmidt *et al.* procedures.

Raju, Cabrera and Lezotte model. Both the BCG and RBN models are based on the implicit assumption that employee performance ratings are continuous. Bobko, Karren and Kerkar (1987) argued that this assumption is not always tenable. More often than not employee performance ratings are made using categorical scales such as 1 to 5. In response to this criticism, Raju, Cabrera and Lezotte (1996) developed a new utility model (RCL) in which performance is viewed as categorical rather than continuous. A special case of this model is the dichotomous utility model in which the number of categories is limited to two. For example, the two categories could be meeting requirements/not meeting requirements of the job or successful/unsuccessful on the job.

The RCL model is based on logistic regression, which is recommended over linear regression when one variable is continuous, X , and the other is dichotomous, Y (Agresti 1984;

Hosmer and Lemeshow 1989). According to the model, P_1 is the probability that an employee with a given predictor score, X , belongs to the successful group and may be expressed as:

$$P_1 = P(1|X) = \frac{e^{aX+b}}{1 + e^{aX+b}} \quad (7)$$

where a and b are the multiplicative and additive factors, respectively, to be determined, and e is a mathematical constant. The probability that a person with a given X belongs to the unsuccessful group, P_0 , may be written as:

$$P_0 = P(0|X) = 1 - P_1 = \frac{1}{1 + e^{aX+b}} \quad (8)$$

The average gain in utility of a new selection procedure (II) over an old procedure (I) may be expressed as:

$$\Delta \bar{U} = (Y_1 - Y_0)(\bar{P}_{(II)1} - \bar{P}_{(I)1}) \quad (9)$$

where Y_1 is the estimated monetary value of an average successful employee, Y_0 is that of an average unsuccessful employee and $\bar{P}_{(II)1}$ and $\bar{P}_{(I)1}$ are the average probabilities of success (given selection) for the new and old selection procedures, respectively.

Like the RBN model, the RCL model does not require the difficult estimation of SD_Y . However, it does require that supervisors estimate the monetary value of an average successful employee and of an average unsuccessful employee. These estimates may be easier for supervisors to provide than the percentile estimates required for the Schmidt *et al.* procedure because they more closely mirror supervisors' perceptions of employee performance. For example, Karren and Bobko (1983) found that some supervisors were unwilling or unable to make distinctions between employees at the 50th and the 85th percentiles. They considered that employees were either meeting requirements (successful) or not meeting requirements (unsuccessful). In cases where supervisors perceive employee performance to be categorical, the RCL model may be appropriate.

Current Trends in UA Research

UA Acceptance

Much recent UA research has shifted the focus of study from the measures to issues of acceptance of utility recommendations (Carson, Becker and Henderson 1998; Cronshaw 1997; Hazer and Highhouse 1997; Latham and Whyte 1994; Whyte and Latham 1997). Some of the findings have been disappointing for proponents of UA. A study by Latham and Whyte (1994) revealed that the presentation of utility data

regarding a proposed selection procedure actually reduced manager support for the program. That is, managers were less likely to support the implementation of a new selection procedure when given information about the utility of the program than when utility information was not presented. Managers reacted more favorably when presented with information regarding the validity of the selection procedure or validity plus expectancy table information.

Whyte and Latham (1997) hypothesized that, given the positive impact of source credibility on message believability, the negative findings of their previous study could possibly be eliminated if an expert explained UA to the managers. A replication of their own study including a new scenario in which an expert explained UA and was available to answer questions again found that managers were less likely to support a new selection program when utility information was provided.

An interesting commentary on Whyte and Latham's study was made by Cronshaw (1997), who acted as the expert in their study. Cronshaw suggested that the utility information that he presented may well have been perceived as more of a high pressure or coercive communication rather than a neutral message intended to inform. He proposed that managers' negative reactions to this attempt to 'sell' the selection intervention could have caused the results. Cronshaw concluded that managers are more likely to accept UA as a decision tool when it serves an informational rather than a persuasive purpose.

In fact, not all studies of UA acceptance have resulted in negative reactions. Hazer and Highhouse (1997) presented managers with different UA scenarios and found that their perceptions of the usefulness of UA were above the midpoint for all scenarios. They also found that managers had more favorable reactions toward UA when the 40% method was used to estimate SD_{γ} . In another replication of Latham and Whyte's study, Beckstein and Gilliland (1996) used exactly the same materials and failed to find lower acceptance when UA information was presented. In a second study they revised the UA materials to make them more 'user friendly' and found that this information was more favorably viewed than validity information alone or validity plus original utility information.

This conclusion was corroborated by Carson, Becker and Henderson (1998) who replicated Latham and Whyte's original study using MBA students. They included an additional scenario in which utility information was presented in a more understandable way. Results revealed low to moderate positive effects for this type of utility information on levels of acceptance. They

concluded that the manner in which utility information is presented could affect the potential acceptance of UA recommendations.

These studies raise a number of issues to be investigated regarding the acceptance of UA recommendations. Most of the conclusions have suggested that the way in which UA information is presented is an important factor for acceptance. More 'user friendly' explanations of UA have resulted in more positive reactions. Also, presenting UA in an informational rather than a persuasive manner may also help to increase acceptance of UA recommendations. In concluding, it should be noted that, while several studies have failed to replicate the negative UA reactions found by Latham and Whyte, none of them has found high levels of support by managers for UA. At best low to moderate levels of acceptance have been encountered.

Multi Attribute UA

Another possible factor influencing the acceptance of UA may be that an estimate of the monetary value of selection programs is not always the only information that managers are interested in when making their decisions. Macan and Highhouse (1994) reported that managers often do not request information regarding the monetary value of HR activities. Multi attribute utility (MAU) models are utility models which incorporate multiple outcomes or attributes, not just the value of job performance, into the decision making process. MAU allows decision makers to evaluate a number of decision options by considering the effects of each option on various attributes or outcomes that may be important to the organization (Boudreau 1991).

MAU analysis requires that decision makers first make a list of attributes that they consider important for making the final choice among selection procedures. Roth and Bobko (1997) present an example of a selection system for which relevant attributes include diversity, legal exposure and organizational image in addition to increased value of job performance. Each of the chosen attributes is measured and combined into a single composite number that represents the benefit for each of the possible selection procedures.

There are often multiple factors that influence organizational decisions. Kaplan and Norton (1996) illustrate this point by describing a cockpit of a jet airplane that has only one instrument measuring 'airspeed'. They question how the reader would feel about flying on that plane. Even if the pilot did an excellent job on airspeed, how could he avoid colliding with tall mountains or running out of fuel? Clearly pilots need information from a large set of indicators to navigate an airplane. Organizational decision

makers also need to consider more than one piece of information when choosing among alternatives. While the value of performance in monetary units will surely continue to be one of the most important attributes in any decision making process within organizations, MAU models offer decision makers the possibility of including other important outcomes in their utility equation.

Decision Making Processes

The aforementioned studies regarding UA acceptance have made the need to better understand the decision making processes of managers another area of interest for UA researchers. Boudreau pointed out that, 'UA models offer normative theories about the factors that decision makers should consider in making HRM decisions, but actual decisions probably depart from UA prescriptions' (1991, p. 710).

A promising advance toward understanding the decision making process as it relates to UA acceptance comes from the work of Roth, Segars and Wright (2000). They define the domain of UA acceptance and then present a theoretical perspective that helps to delineate constructs and suggest future research directions. An important distinction the authors make is between the 'pre use domain' of UA acceptance, in which decision makers decide whether or not to conduct a UA, and the 'post use domain', in which decision makers choose whether or not to accept the recommendations once UA has been implemented. The acceptance studies up to date have clearly focused on the post use domain. Finally, Roth *et al.* (2000) use Image Theory, a behavioral theory of decision making in organizations developed by Beach and Mitchell (1996), to analyse the process decision makers go through when deciding whether or not to conduct a UA. They then draw on dual process theories from the persuasion literature to analyse post use decisions to accept UA results.

Participation

Many researchers have argued that allowing decision makers to participate in the utility analysis process will increase acceptance of UA results. Roth *et al.* (2000) mentioned the positive effect that participation has been found to have on the acceptance of numerous organizational interventions. Participation may lead to higher levels of acceptance because individuals have greater knowledge of the issue under consideration and/or because commitment to an issue tends to be greater when individuals participate in its planning or design. Likewise, Roth and Bobko (1997) suggested that a potential

advantage of MAU models is that they require the participation of decision makers in choosing, measuring and weighing the relevant attributes. They stated that this participation would likely increase acceptance of the final recommendations. Rauschenberger and Schmidt (1987, p. 55) proposed that one way to increase the perceived credibility of UA results was to 'get organizational decision makers involved in the utility analysis at the earliest possible time.'

Finally, Morrow, Jarrett and Rupinski (1997) reported a positive effect of participation on the acceptance of UA results in their four year investigation of the utility of one organization's corporate wide training. In their case, a multi attribute utility model based on the RBN model was modified and approved by managers in a series of meetings previous to the application of the model. The authors concluded that managers accepted the utility estimates due to managerial pre approval of the utility model and because the managers were 'able to identify and control the parameters that were included in the multi attribute utility model' (Morrow *et al.* 1997, p. 114).

Strategic UA

The instigation of the previously mentioned utility investigation conducted by Morrow *et al.* (1997) was a request by the organization's CEO for a report of the monetary value of training. The CEO indicated that training should be aligned with the organization's strategic goals and should be experimentally evaluated in order to show that it was a worthwhile investment. There is a clear trend among organizations to question the strategic contribution of their human resource interventions (Huselid, Jackson and Schuler 1997; Ulrich 1997; Wright and McMahan 1992). In response to this demand, Cabrera and Cabrera (1999) proposed a strategic utility analysis model. Using the MAU framework, which considers multiple outcomes in the analysis of utility, the authors suggest that the specific outcomes included should come directly from the management control framework an organization uses to assess its strategic performance. In this way UA methods incorporate the same indicators of success that the rest of the organization uses. These indicators may include anything from customer satisfaction to sales growth to production errors. This kind of UA procedure will provide decision makers with information regarding the relative contributions of different selection alternatives toward the achievement of the organization's strategic objectives.

Future Directions for UA Research

Accuracy of UA Estimates

While there are many models for assessing the utility of a selection program or an organizational intervention, there is very little reported research on the accuracy of utility analysis estimates. Alexander and Barrick (1987) proposed some approximate standard errors for utility estimates, and Anderson and Muchinsky (1991) and Quartetti and Raju (1998) reported some Monte Carlo results on the distribution of utility estimates.

The results of Quartetti and Raju's (1998) study showed that sample size did not affect average utility estimates, whereas utility estimates were found to increase as both the population validity increases and as the predictor distribution becomes more negatively skewed. Utility estimates also increased as the selection ratio decreased. Finally, for all of the analyses, estimates based on the RCL model were more accurate and had less variability than those based on the BCG model.

While these studies are important, more research is still needed to develop appropriate standard errors of various utility estimates. Without some indication of the degree of accuracy associated with the currently popular utility estimates, it may be difficult to take such estimates very seriously.

Accuracy of SD_Y Estimation

Although several empirical studies have compared the Schmidt *et al.* (1979) procedure with the CREPID procedure in estimating SD_Y , the accuracy of these estimates is still an unanswered question. Based on these empirical studies, we appear to have gained a general understanding of how the different estimates of SD_Y compare with one another, but no such understanding is currently available regarding the accuracy of the SD_Y estimates. One study regarding the robustness of SD_Y estimates (Anderson and Muchinsky 1991) did find that SD_Y estimates using Schmidt *et al.*'s (1979) percentile estimates were fairly robust; that is they were not strongly affected by departures from normality.

Thus, there is a definite need for a comprehensive assessment of the reliability and validity/accuracy of the SD_Y estimates. The number of managers typically involved in estimating SD_Y is rather small in both the Schmidt *et al.* (1979) and CREPID procedures. At the present time, very little is known about the effect of these small sample sizes on the estimation of SD_Y . Bobko *et al.* (1987) have previously emphasized the need for such an empirical investigation.

UA Acceptance

As noted earlier, the acceptability of UA results by managers should not be taken for granted. UA researchers have attributed low levels of acceptance by managers and other consumers of UA results to a lack of understanding of the UA framework, its assumptions, and its terminology. UA proponents have argued for a need to adopt the commonly accepted business accounting terminology in describing UA results. Most also agree that acceptance levels can be increased by getting managers involved upfront in the formulation of appropriate, company wide strategies for the development and implementation of multi attribute organizational interventions and the assessment of their utility.

But perhaps getting managers to better understand UA is not the only problem. UA researchers should also attempt to better understand decision makers' needs. Rather than trying to 'sell' UA methods to managers, UA proponents should first ask them what information they want when they are making decisions. Maybe UA procedures are not accepted because they do not offer managers what they want. Remember that Macan and Highhouse (1994) found that many HR managers are not interested in the monetary value of their interventions. In fact, there is a general trend in the area of management control towards incorporating multiple, non financial indicators of success when assessing organization wide performance (Kaplan and Norton 1992).

Current UA methods and procedures may be extremely useful if they are adapted to provide decision makers with the information that they consider to be important. MAU methods are a step in this direction. Why not use UA procedures to show managers which selection procedure will provide the greatest increase in customer satisfaction or the greatest reduction in errors, if that is what they want to know?

A Comparison of Models

While there is a great deal of empirical data concerning the BCG model, there is only limited information on how the new models (RBN and RCL) compare with the BCG model in terms of their utility estimates. There is a need for systematic investigation of the various models (including the BCG extensions proposed by Boudreau and Berger (1985a, 1985b) and De Corte (1994, 1996, 1998)) with respect to implementation problems and the congruence of utility estimates. Such investigations are essential if we are to take full advantage of the various utility models and for generating greater acceptance of the utility models among managers and business executives.

Conclusion

Organizational decision makers need decision aids to help them evaluate the relative contributions of different selection programs. Utility analysis models offer concrete procedures for evaluating the monetary impact of selection options. Extensions of the traditional UA models and the development of new models have made UA more readily applicable for practitioners. Given that managers are now faced with a number of alternative UA methods, it is important that these alternatives be further investigated to provide clearer evidence regarding the accuracy of their respective estimates.

Another key issue for UA research is that of acceptance. Involving managers in the utility analysis process appears to play a vital role in UA acceptance. Efforts are being made to better understand the managerial decision making processes that influence when managers choose to apply UA models and under what circumstances they are more likely to accept the final UA recommendations. Again, maybe a more fruitful direction for UA research would be to determine what information managers want when they are making decisions and to incorporate this information into current UA models. By allowing for the inclusion of a number of outcomes in the analysis of alternatives, multi attribute utility models are a step in this direction.

In summary, while UA researchers have come far in developing practical, realistic procedures for assessing the relative impact of diverse selection procedures, there is still much to be done. Issues of applicability, accuracy and acceptance abound. There continue to be opportunities for research in comparing different models and testing the accuracy of SD_Y and utility estimates. Meanwhile there is exciting new ground to be covered regarding how to increase the use and acceptance of utility models by managers or possibly how to reorient current UA models so that they provide decision makers with the information they need and want.

Acknowledgement

This research was funded in part by the Spanish Ministry of Science and Technology SEC 2000 0395.

References

- Agresti, A. (1984) *Analysis of Ordinal Categorical Data*. New York: John Wiley and Sons.
- Alexander, R.A. and Barrick, M.R. (1987) Estimating the standard error of projected dollar gains in utility analysis. *Journal of Applied Psychology*, **72**, 475-479.
- Anderson, S.M. and Muchinsky, P.M. (1991) An examination of the robustness of the general utility function. *Educational and Psychological Measurement*, **51**, 49-65.
- Beach, L.R. and Mitchell, T.R. (1996) Image theory, the unifying perspective. In L.R. Beach (ed.), *Decision Making in the Workplace: A Unified Perspective* (pp. 1-20). Mahwah, NJ: Erlbaum.
- Beckstein, B.A. and Gilliland, S.W. (1996) The applicability of utility analysis: When do utility estimates influence decisions? Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA, April.
- Bobko, P., Karren, R. and Kerkar, S.P. (1987) Systematic research needs for understanding supervisor based estimates of SD_Y in utility analysis. *Organizational Behavior and Human Decision Processes*, **40**, 69-95.
- Bobko, P., Karren, R. and Parkington, J.J. (1983) Estimation of standard deviations in utility analyses: An empirical test. *Journal of Applied Psychology*, **68**, 170-176.
- Boudreau, J.W. (1983a) Economic considerations in estimating the utility of human resource productivity improvement programs. *Personnel Psychology*, **36**, 551-576.
- Boudreau, J.W. (1983b) Effects of employee flows on utility analysis of human resource productivity improvement programs. *Journal of Applied Psychology*, **68**, 396-406.
- Boudreau, J.W. (1991) Utility analysis for decisions in human resource management. In M.D. Dunnette and L.M. Hough (eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 2, pp. 621-745). Palo Alto, CA: Consulting Psychologists Press.
- Boudreau, J.W. and Berger, C.J. (1985) Decision theoretic utility analysis applied to external employee movement. *Journal of Applied Psychology*, **70**, 581-612.
- Brogden, H.E. (1949) When testing pays off. *Personnel Psychology*, **2**, 171-185.
- Burke, M.J. and Frederick, J.T. (1986) A comparison of economic utility estimates for alternative SD_Y estimation procedures. *Journal of Applied Psychology*, **71**, 334-339.
- Burke, M.J., Raju, N.S., Edwards, J.E. and Day, R.R. (1993) Performance scale transformations in utility analysis. Unpublished manuscript, Tulane University, New Orleans.
- Cabrera, E.F. (1998) Análisis de utilidad: La valoración del impacto financiero de los programas de recursos humanos. *Psicológica*, **19**, 217-236.
- Cabrera, E.F. and Cabrera, A. (1999) Strategic Utility Analysis. Paper presented at the Annual Meeting of the Academy of Management, Chicago, IL, August.
- Carson, K.P., Becker, J.S. and Henderson, J.A. (1998) Is utility really futile? A failure to replicate and an extension. *Journal of Applied Psychology*, **83**, 84-96.
- Cascio, W.F. (1991) *Costing Human Resources: The Financial Impact of Behavior in Organizations*. Boston: PWS Kent Publishing Company.
- Cascio, W.F. and Ramos, R.A. (1986) Development and application of a new method for assessing job

- performance in behavioral/economic terms. *Journal of Applied Psychology*, **71**, 20-28.
- Cronbach, L.J. and Gleser, G.C. (1965) *Psychological Tests and Personnel Decisions*. Urbana, IL: University of Illinois Press.
- Cronshaw, S.F. (1997) Lo! The stimulus speaks: The insider's view on Whyte and Latham's 'The futility of utility analysis'. *Personnel Psychology*, **50**, 611-615.
- De Corte, W. (1994) Utility analysis for the one cohort selection retention decision with a probationary period. *Journal of Applied Psychology*, **79**, 402-411.
- De Corte, W. (1996) Recruitment and retention decisions that maximize the utility of probationary selection to obtain a fixed quota of successful selectees. *Personnel Psychology*, **49**, 399-428.
- De Corte, W. (1998) Estimating and maximizing the utility of sequential selection decisions with a probationary period. *British Journal of Mathematical and Statistical Psychology*, **51**, 101-121.
- Hazer, J.T. and Highhouse, S. (1997) Factors influencing managers' reactions to utility analysis: Effects of SD_Y method, information frame, and focal intervention. *Journal of Applied Psychology*, **82**, 104-112.
- Hosmer, D.W. and Lemeshow, S. (1989) *Applied Logistic Regression*. New York: John Wiley and Sons.
- Hunter, J.E. and Schmidt, F.L. (1982) Fitting people to jobs: The impact of personnel selection on national productivity. In M.D. Dunnette and E.A. Fleishman (eds.), *Human Performance and Productivity: Human Capacity Assessment* (Vol. 1, pp. 232-284). Hillsdale, NJ: Erlbaum.
- Hunter, J.E., Schmidt, F.L. and Judiesch, M.K. (1990) Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*, **75**, 28-43.
- Huselid, M.A., Jackson, S.E. and Schuler, R.S. (1997) Technical and strategic human resource management effectiveness as determinants of firm performance. *Academy of Management Journal*, **39**, 949-969.
- Kaplan, R.S. and Norton, D.P. (1992) The Balanced Scorecard measures that drive performance. *Harvard Business Review*, **January/February**, 71-79.
- Kaplan, R.S. and Norton, D.P. (1996) Linking the Balanced Scorecard to strategy. *California Management Review*, **39**, 53-79.
- Karren, R. and Bobko, P. (1983) Conducting utility analyses: Some methodological concerns. Paper presented at the Annual Meeting of the Academy of Management, Dallas, TX, August.
- Latham, G. and Whyte, G. (1994) The futility of utility analysis. *Personnel Psychology*, **47**, 31-46.
- Law, K.S. and Myors, B. (1993) Cutoff scores that maximize the total utility of a selection program: Comment on Martin and Raju's procedure. *Journal of Applied Psychology*, **78**, 736-740.
- Lezotte, D.V., Raju, N.S., Burke, M.J. and Normand, N. (1996) An empirical comparison of two utility analysis models. *Journal of Human Resource Costing and Accounting*, **1**, 19-30.
- Macan, T.H. and Highhouse, S.H. (1994) Communicating the utility of human resource activities: A survey of I/O and HR professionals. *Journal of Business and Psychology*, **8**, 425-436.
- Martin, S.L. and Raju, N.S. (1992) Determining cutoff scores that optimize utility: A recognition of recruiting costs. *Journal of Applied Psychology*, **77**, 15-23.
- Morrow, C.C., Jarrett, M.Q. and Rupinski, M.T. (1997) An investigation of the effect and economic utility of corporate wide training. *Personnel Psychology*, **50**, 91-119.
- Murphy, K.M. (1986) When your top choice turns you down: The effect of rejected offers on the utility of selection tests. *Psychological Bulletin*, **99**, 133-138.
- Quartetti, D.A. and Raju, N.S. (1998) A Monte Carlo assessment of estimation in utility analysis. Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, Dallas, Texas, April.
- Raju, N.S., Burke, M.J. and Normand, J. (1990) A new approach for utility analysis. *Journal of Applied Psychology*, **75**, 3-12.
- Raju, N.S., Cabrera, E.F. and Lezotte, D.V. (1996) Utility analysis when employee performance is classified into two categories: An application of three utility models. Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, San Diego, CA, April.
- Rauschenberger, J.M. and Schmidt, F.L. (1987) Measuring the economic impact of human resource programs. *Journal of Business and Psychology*, **2**, 50-59.
- Reilly, R.R. and Smither, J.W. (1985) An examination of two alternative techniques to estimate the standard deviation of job performance in dollars. *Journal of Applied Psychology*, **70**, 651-661.
- Roth, P.L. and Bobko, P. (1997) A research agenda for multi attribute utility analysis in human resource management. *Human Resource Management Review*, **7**, 341-368.
- Roth, P.L., Segars, A.H. and Wright, P.M. (2000) Utility analysis acceptance research: A review and analysis. Unpublished manuscript, Clemson University.
- Sands, W.A. (1973) A method for evaluating alternative recruiting selection strategies: The CAPER model. *Journal of Applied Psychology*, **57**, 222-227.
- Schmidt, F.L., Hunter, J.E., McKenzie, R.C. and Muldrow, T. (1979) The impact of valid selection procedures on workforce productivity. *Journal of Applied Psychology*, **64**, 609-624.
- Taylor, H.C. and Russell, J.T. (1939) The relationship of validity coefficients to the practical effectiveness of tests in selection. *Journal of Applied Psychology*, **23**, 565-578.
- Ulrich, D. (1997) *Human Resource Champions: The Next Agenda for Adding Value and Delivering Results*. Boston: Harvard Business School Press.
- Weekley, J.A., Frank, B., O'Connor, E.J. and Peters, L.H. (1985) A comparison of three methods of estimating the standard deviation of performance in dollars. *Journal of Applied Psychology*, **70**, 122-126.
- Whyte, G. and Latham, G. (1997) The futility of utility analysis revisited: When even an expert fails.