

THE FUTILITY OF UTILITY ANALYSIS REVISITED: WHEN EVEN AN EXPERT FAILS

GLEN WHYTE, GARY LATHAM
Rotman School of Management
University of Toronto

The effect of utility analysis on managerial decisions regarding the use of a valid selection test was investigated. Experienced managers ($n = 41$) were randomly assigned to one of three conditions. The control condition contained written advice to adopt new selection procedures from a hypothetical psychologist that included an explanation of validation procedures. The second condition contained the same written information and advice as the control condition plus written support of that advice from a hypothetical trusted adviser. The third condition contained all the material that was presented in the control condition plus a written explanation of utility analysis, an actual utility analysis showing that large financial benefits would flow from using the proposed procedures, and a video-taped presentation from an expert on utility analysis where the logic underlying utility analysis and its benefits were explained. The expert was then presented live to the audience to address any questions that might have arisen from the utility analysis or the video. An ANOVA revealed that the presentation of a positive utility analysis reduced support for implementing a valid selection procedure, even though the logic and merits of utility analysis as a decision-making tool were described by an internationally recognized authority.

Many I/O psychologists consider it useful to place a relatively accurate monetary value on the contribution of interventions based on I/O psychology to an organization's effectiveness. One technique, utility analysis, is used to forecast the net financial benefits of human resource management (HRM) initiatives. Utility analysis has been applied to selection procedures (e.g., Cronshaw, 1986, 1991; Cronshaw & Alexander, 1985), performance appraisal (Florin-Thuma & Boudreau, 1987), training (Cascio, 1991; Schmidt, Hunter, & Pearlman, 1982), promotion (Cascio & Ramos, 1986), recruitment (Boudreau & Rynes, 1985), and turnover/downsizing (Boudreau & Berger, 1985; Cascio, 1991).

This research was supported in part by a grant awarded to the second author from the Social Sciences and Humanities Research Council of Canada

Correspondence and requests for reprints should be addressed to Professor G. Whyte, University of Toronto, Rotman School of Management, 105 St. George St., Toronto, Ontario, Canada M5S 3E6, e-mail: whyte@fmgmt.mgmt.utoronto.ca

The assumption that managers find the information provided by utility analysis useful when making HRM decisions is questionable. Latham and Whyte (1994) found that experienced managers were indeed influenced by utility analysis, but not in the way that was intended by advocates of this technique. A utility analysis indicating that substantial net benefits would flow if a valid selection procedure was implemented actually decreased managerial support for the selection test. These findings were surprising, and undercut a central assumption of utility analysis that managers find this type of information helpful in decision making.

In this study, we attempted to replicate and extend the findings of Latham and Whyte (1994): The rationale for this extension is two-fold. The first is ecological validity. In many cases, when utility analysis is employed there is an opportunity for the individual using it to explain the technique and its merits. A possible criticism of Latham and Whyte's study is that the utility analysis was explained in writing to managers. No expert in utility analysis was available as an advocate of the technique. Had an expert been present to explain the technique, it might have been accepted by managers.

The second rationale is based on research in social psychology that has led to the development of general models of social persuasion and attitude change (e.g., Hovland, Janis, & Kelley, 1953; McGuire, 1985; Petty & Cacioppo, 1986). This literature provides a useful perspective from which to consider the effect of utility analysis on managerial judgment and decision making. Studies of the persuasion process demonstrate that characteristics of the source of a communication can determine the success or failure of attempts to influence the attitudes and behaviors of others (McGuire, 1985). A general finding is that source credibility based on expertise affects the believability of a persuasive communication (Hovland & Weiss, 1951; Petty & Cacioppo, 1986). An expert in utility analysis might therefore have much more success in convincing managers to use the results of utility analysis in decision making than did the hypothetical psychologist used by Latham and Whyte (1994).

Our interest was in whether the findings of Latham and Whyte (1994) would extend to the situation in which an expert on the topic of utility analysis explained the value of this information to managers. The findings of Latham and Whyte (1994) should be considered robust only if a live expert on utility analysis, who was neutral as to the decision taken, would be unable to convince managers to rely on utility analysis to assist them in deciding whether to implement a valid selection procedure.

Method

Sample

Forty-one managers, 10 female and 31 male, participated in the study. The managers had an average of 14 years ($SD = 6.6$) full-time work experience. The average age of the managers was 37 years old ($SD = 6.2$). All managers were enrolled in an Executive MBA (EMBA) program at a North American University. None of them had been exposed to utility analysis prior to the study.

Study Design

This study employed a randomized, single-factor design. The factor is called Information, and refers to whether managers were provided with information regarding either validity, validity plus the written opinion of a trusted adviser, or validity plus a written presentation of utility analysis and a video-tape on utility analysis featuring an actual expert, followed by the expert making himself available to the managers to answer their questions. Information is thus a three-level factor.

There were three unique sets of stimulus materials used in this study. All materials contained instructions, a description of a decision scenario, and one of three information conditions. Each manager received only one of the three versions of the stimulus materials. The materials were randomly distributed to participants, subject to the constraint that each of the three conditions contain approximately the same number of participants.

Procedure

All managers read a scenario in which they were the vice-president of a large multi-national corporation that was experiencing problems in the quality of recently hired clerical/administrative personnel. A highly qualified and experienced psychologist had been retained to investigate this issue. The psychologist recommended in writing the adoption of new selection practices that were more valid than existing practices and hence would improve the quality of the employees who are hired. The managers were told that they had the authority to decide whether to follow the psychologist's advice.

In the control condition, an accomplished hypothetical psychologist advises in writing that the implementation of valid selection practices will improve employee performance, and supports that advice with a written explanation of standard validation procedures. This condition was based

on a frequent occurrence in organizations, namely, the presentation to management of a report written by a psychologist that explains the concept of validity and recommends the use of a selection test. This condition is identical to the control condition used by Latham and Whyte (1994).

The first experimental condition included all information contained in the control condition, plus an additional paragraph. In this paragraph, the managers were told that a hypothetical trusted adviser, who was described as experienced and "street-smart," estimated that the gain would be substantial if the psychologist's advice were followed. No mention of utility analysis was made. This condition allowed us to examine the extent to which managers find the opinion of a trusted adviser persuasive in deciding whether to implement a valid selection procedure. If managers are not swayed by quantitative analysis and hard data, then according to Mintzberg (1975; 1989) they may be swayed by the opinion of someone whom they trust.

The second experimental condition contained all the written information given to participants in the control condition, an additional written section in which utility analysis was explained, and a written description of a utility analysis that the psychologist had performed that supported the proposed changes in selection procedures. The psychologist provided an estimate of the approximate pre-tax financial gain achievable if the new selection practices were implemented. She explained that the model used to calculate this amount was developed and refined over many years of practice, and that it is used by most I/O psychologists. Each term in the utility equation was defined. In short, this condition included everything the authors could conceive of to ensure that utility analysis would be understood and accepted by the managers. This condition was based on that used by Latham and Whyte (1994).

Prior to reading this material, the managers in this condition watched a professionally recorded video entitled "Utility Analysis with Dr. Steven Cronshaw." In this video, Dr. Cronshaw is introduced as a Professor of Psychology and Past Chair of the Industrial-Organizational Psychology Division of the Canadian Psychological Association, and as an expert in utility analysis with considerable experience as a consulting psychologist in industry.

Dr. Cronshaw communicated cogently on video-tape that utility analysis is a valid and useful technique for forecasting the dollar value of human resource management interventions. Dr. Cronshaw stressed that utility analysis should be used in making HRM decisions because (a) it is well researched, field tested, and very sound; (b) many organizations have relied on it for a long time; and (c) the basic idea underlying utility analysis is straightforward and logical. The video concludes with

Dr. Cronshaw strongly recommending, as both a psychologist and a former accountant, the use of utility analysis because it helps managers to make sound decisions.

A videotape of the presentation was used because videotapes are an implicit signal in the EMBA classroom that the presented material is well documented. Moreover, this methodology allowed Dr. Cronshaw to prepare and refine a presentation to his satisfaction that would persuade managers to use utility analysis in their decision making. After the managers saw the video, Dr. Cronshaw in person requested the managers to ask him any questions they wished regarding utility analysis in general, or the specific utility analysis that had just been presented to them. Because of the results obtained by Latham and Whyte (1994), source of expertise and amount of information were deliberately confounded so as to increase the probability that the recommendation based on utility analysis would be accepted.

Measures

After reading the scenario, the managers in all three conditions responded to a 9-item decision preference scale. This scale is identical to that used by Latham and Whyte (1994), with the exception of one additional item. The scale was designed to probe managers' preferences regarding the decision to implement the psychologist's recommendations. The scale assessed such issues as commitment to implementing the psychologist's recommendations, confidence in the ability of the psychologist's advice to improve the quality of new hires, perceptions of the quality of the psychologist's advice, perceptions of the size of the gain that would follow from implementing the psychologist's recommendations, and perceptions of managers' ability to justify to others their decision whether to heed the psychologist's advice.

The managers responded to the items using a 5-point Likert-type scale. A response of "1" indicated a preference to discard the consultant's advice, whereas a response of "5" indicated a preference to follow the advice. The reliability coefficient (Cronbach's alpha) for the entire scale was .93. A factor analysis relying on the method of principal component analysis for factor extraction confirmed the unitary structure of the scale.

The statistical analyses examined the effects of the information manipulations on managers' willingness to follow the psychologist's advice. The overall effects of the manipulations were tested, followed by planned pairwise comparisons between means.

TABLE 1
Decision Preference Scale Score—Means and Standard Deviations

Information condition	Score (SD)
Control	32.8 (6.3)
Trusted adviser	33.9 (5.3)
Expert utility analysis	26.5 (9.3)

Results

Decision preference scale scores could range from 9–45, with a high number indicating a strong preference to adopt the psychologist's recommendation. The cell means and standard deviations for decision preference scale score are presented in Table 1 for the three-level single-factor design. A one-way analysis of variance was conducted on managers' responses to the decision preference scale. The results indicate that information had a statistically significant main effect on willingness to adopt the psychologist's advice ($F(2,39) = 4.15, p < .02$).

Planned comparisons were conducted next. As shown in Table 2, the difference between the control condition and the trusted adviser condition was not statistically significant.

In contrast, the difference between the live expert utility analysis condition and the control condition was significant. The level of acceptance of the selection test was higher when the recommendation was based solely on a written report from a hypothetical psychologist where the concept of validity was explained, than it was when this same information plus a presentation on utility analysis from an actual psychologist was given to managers.

The results of the comparison between the live expert utility analysis condition and the trusted adviser condition were also statistically significant. Managers were more inclined to use the selection test when it was recommended by a hypothetical trusted adviser than when its use was supported by a utility analysis and a presentation from an actual psychologist who is an expert on utility analysis.

The standardized effect sizes calculated for each comparison are shown in Table 2. The effect sizes of the statistically significant comparisons between the utility analysis and both the control and trusted adviser conditions respectively, were .90 and 1.06. According to convention, the operational definition of small, medium, and large differences between means corresponds to standardized effect sizes of .2, .5, and .8, respectively (Cohen, 1977) The effect size associated with the drop in support for the psychologist's advice caused by an actual expert's presentation of utility analysis is therefore large.

TABLE 2
Decision Preference Scale Score—Results of Pairwise Comparisons

Comparisons	Difference	Effect size index	Significance
A-C	6.3	.90	.05
B-C	7.4	1.06	.02
A-B	-1.1	.16	<i>ns</i>

A-control

B-trusted adviser

C-expert utility analysis

ns-non-significant

To understand why managers were least inclined to follow the psychologist's advice in the utility condition, we examined responses to the decision preference scale on an item by item basis. Responses to all nine scale items were consistent with the view that managers were less convinced of the wisdom of the psychologist's advice in the utility condition than the control condition. The differences were greatest and were statistically significant for items 5 (How committed are you to implementing the consultant's recommendations?), 8 (How effectively could you justify to others in the firm a decision to implement the consultant's recommendations?), and 9 (How important in your decision making was the estimate of the financial consequences of the consultant's recommendations?); $F(2,39) = 3.9, p < .03$; $F(2,39) = 4.65, p < .02$; $F(2,39) = 5.55, p < .01$, respectively.

To investigate whether the use of the pronoun "she" to describe the psychologist had an unintended effect, we examined whether the responses of the males and females were different. No statistically significant differences were found for the decision preference scale, $F(1,40) = 2.05, p < .16$, or for any of the items that make up the scale.

We also investigated whether those managers with a finance position ($n = 15$) were influenced by the utility analysis differently from those without a finance background. An ANOVA revealed no statistically significant difference between the groups on the decision preference scale, $F(1,40) = .03, p < .86$.

We further investigated whether people ($n = 14$) in human resources (HR), who presumably were familiar with the concept of validity, might have responded differently from those without an HR background. Again, no significant difference in responses to the decision preference scale was found, $F(1,40) = .04, p < .85$.

Discussion

The present results challenge the core assumption of utility analysis, namely, that managers find it useful in deciding whether to invest in human resource management systems such as selection tests. Not only did the results of a highly positive utility analysis fail to convince managers of the need for a change in HR policy, the presentation of the analysis also undercut support for such a change. These findings replicate, under more stringent conditions, those of Latham and Whyte (1994). First, an actual expert in utility analysis explained and actively tried to persuade experienced managers to use the results of a utility analysis to make a positive decision. Even an expert, however, could not persuade managers to overcome their reluctance to rely on the information provided by a utility analysis.

Second, the managers had the opportunity to question the expert in person on any aspect of the analysis prior to responding to the decision preference scale. None chose to do so.

Third, all managers were enrolled in an EMBA program, and as such were both trained in and experienced with applying quantitative and other forms of rational analysis to managerial problems. And yet the presentation of the utility analysis still reduced support for the recommendation to use a valid selection test.

Managers' lack of commitment to the psychologist's advice in the utility condition appears to be most strongly linked to the perception that a utility estimate would be unhelpful in attempting to sell the psychologist's advice to others in the firm, and that this information is unimportant in the decision whether to proceed with the psychologist's advice.

The present results do not impugn the scientific validity or predictive accuracy of utility analysis. Nonetheless, the challenge posed by the results is clear and important. That challenge is to the utility of utility analysis as perceived by experienced managers. These results do not support the use of utility analysis as a tool to assist managerial decision making. I/O psychologists are advised on the basis of these findings to reconsider their assumptions regarding the information managers value when making HR policy decisions.

Our findings are in some respects confirmation of Johns' (1993) important critique of the assumptions that ground much of the literature on HR strategy formulation. In that article, Johns (1993; pp. 569) suggested that the adoption of I/O type HR practices is "not strongly influenced by technical merit." Bamberger and Fiegenbaum (1996) endorse this viewpoint in their description of a theory that explains how technical, rational-economic, institutional, and political forces combine to help shape how managers decide which HR practices to adopt.

This study has attempted to add to existing knowledge about utility analysis by examining managers' receptivity to this technique when it is advocated and explained by an expert. Further research, however, would be appropriate to address the limitations of the current study. For example, the trusted adviser manipulation could have been stronger. Trust, however, is difficult to manipulate in a laboratory setting. This fact may explain the failure of this manipulation to influence managerial decision making.

Further, no data were collected regarding managers' knowledge of test validation or managers' perceptions of Cronshaw's utility presentation and the utility framework in general. These limitations could affect the conclusions that can be drawn from this research.

More generally, the laboratory methodology employed here raises concerns about external validity. We attempted to minimize those concerns by using experienced managers as participants, as suggested by Gordon, Slade, and Schmidt (1986), and by using an expert to teach the value of utility analysis.

A final prominent limitation of the present research is its low power. A one-way ANOVA using a sample of the size employed in this study and a .05 criterion attains a power greater than the .80 power convention only by assuming a large difference of $h = 1.1$ (Cohen, 1977). This assumption would have been difficult to substantiate at the outset of this research.

The limitations of the present research notwithstanding, the data suggest that psychologists who rely on utility analysis to generate support for I/O type innovation do so at their peril. Psychologists who rely on utility analysis to get their recommendations accepted and succeed in doing so would seem to be successful in spite of, not because of, utility analysis.

REFERENCES

- Bamberger P, Fiegenbaum A. (1996). The role of strategic reference points in explaining the nature and consequences of human resource strategy. *Academy of Management Review*, 21, 926-958.
- 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.
- Cascio WF. (1991). *Costing human resources: The financial impact of behavior in organizations* (3rd ed.). Boston, MS: Kent.
- Cascio WF, Ramos RA. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology*, 71, 20-28.

- Cohen J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cronshaw SF. (1986). The utility of employment testing for clerical/administrative trades in the Canadian military. *Canadian Journal of Administrative Sciences*, 3, 376-385.
- Cronshaw SF. (1991). *Industrial psychology in Canada*. Waterloo, Ontario, Canada: North Waterloo Academic 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.
- Florin-Thuma BC, Boudreau JW. (1987). Performance feedback utility in a small organization: Effects on organizational outcomes and managerial decision processes. *PERSONNEL PSYCHOLOGY*, 40, 693-713.
- Gordon ME, Slade LA, Schmitt N. (1986). The science of the sophomore revisited: From conjecture to empiricism. *Academy of Management Review*, 11, 191-207.
- Hovland CI, Weiss W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15, 635-650.
- Hovland CI, Janis IL, Kelley HH. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.
- Johns G. (1993). Constraints on the adoption of psychology-based personnel practices: Lessons from organizational innovation. *PERSONNEL PSYCHOLOGY*, 46, 569-592.
- Latham GP, Whyte G. (1994). The futility of utility analysis. *PERSONNEL PSYCHOLOGY*, 47, 31-46.
- McGuire WJ. (1985). Attitudes and attitude change. In Lindzey G, Aronson E (Eds.), *The handbook of social psychology*, 2, 233-346. New York: Random House.
- Mintzberg H. (1975). The manager's job: Folklore and fact. *Harvard Business Review*, July-August, 49-61.
- Mintzberg H. (1989). *Mintzberg on management: Inside our strange world of organizations*. Free Press: New York.
- Petty R, Cacioppo J. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Schmidt FL, Hunter JE, Pearlman K. (1982). Assessing the economic impact of personnel programs on workforce productivity. *PERSONNEL PSYCHOLOGY*, 35, 333-347.