Journal of Management
1981, Vol. 7, No. 2, 77-87

Participant Estimates of the Effectiveness of Judgmental Decisions[1]

Stephen A. Stumpf and Dale E. Zand
New York University
Graduate School of Business Administration

Research generally supports the belief that individuals and groups tend to overestimate the effectiveness of their judgmental decisions. There is, however, little information about how consistently these estimates relate to objective measures of effectiveness. Participants' assessments of the anticipated effectiveness of their judgmental decisions were investigated in several studies. It was found that their estimates were poorly and inconsistently related to objective measures of effectiveness. This result raises a need for improved approaches to assessing judgmental decisions and identifying participants who make more accurate judgmental decisions.

Individuals and groups generally tend to overestimate the effectiveness of the solution they have chosen when they have to make a judgmental decision. It has been implicitly assumed that if we can take into account the overestimation bias, then we can use the participant's estimate as a valid indication of the actual effectiveness of the decision.

Unfortunately, little informaton is available about the validity of the relationship between participant estimates and objective measures of decision effectiveness. In infrequent, ill-structured situations such as the Cuban missile crisis or RCA's decision to compete with IBM in computers, estimates of decision effectiveness are crucial and an ineffective decision can be catastrophic. Since objective measures are seldom available until the distant future, a manager often relies on participant estimates of anticipated effectiveness without knowing their validity.

Our purpose is to examine the relation of participant estimates to objective measures of decision effectiveness in ill-structured situations. Since the finding of a low or high relationship in one study would not be a sufficient basis for a reliable conclusion, we sought to replicate our research with different samples, different decisions, and different settings in order to

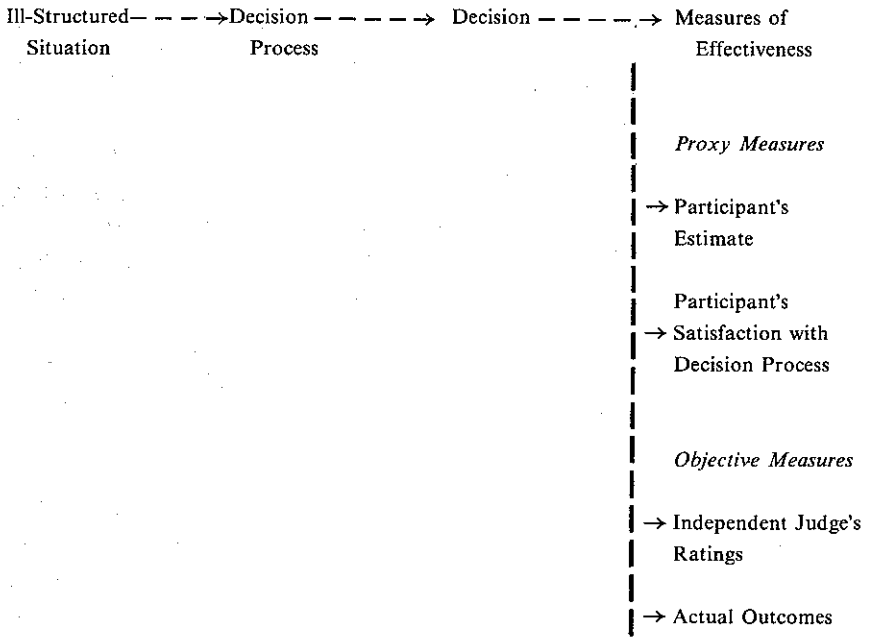
---

Address all correspondence to Stephen A. Stumpf, Associate Professor of Management and Organizational Behavior, New York University, Graduate School of Business Administration, 100 Trinity Place, New York, New York 10006. (212) 285-8938.

[1]Portions of this paper were presented at the American Psychological Association Annual Meeting, Los Angeles, California, August, 1981.

A lenient estimate may or may not be a valid index of actual effectiveness. Validity can be assessed by relating an estimate to operational measures of effectiveness. Whether lenient or not, estimates which lead to accurate inferences are viewed as valid. Cognitive dissonance theory does not address the quality of inferences. It only suggests that estimates of effectiveness will tend to be lenient. In practice, estimates are critical since they are often used to make inferences about effectiveness and thereby affect one's choice.

Figure 1. Assessing the Effectiveness of Judgmental Decisions



This paper is a report of an investigation of the validity of decision makers' estimates of anticipated decision effectiveness that compared them with expert ratings and objective measures of actual outcomes. The research question addressed is: How do participant estimates of the effectiveness of their decision relate to independent ratings and objective measures of decision effectiveness? A related hypothesis is that participants estimate greater effectiveness of their decision than independent ratings and objective measures show them to have.

To examine the above question and hypothesis, three studies were conducted that varied the decision mode and the task: (1) a group decision-making experiment in which business students and managers worked on six different strategic business problems; (2) an individual decision-making simulation with managers making promotion decisions analogous to an actual decision each had made; and (3) a combined group and individual study with managers using a situation with an objective, correct solution.

completely described elsewhere, so we refer the reader to Stumpf, Freedman, and Zand (1979) for more detailed information.

Measures

Three measures were used. Two were subjective, proxy measures and one was an objective, comparison measure.

Group estimate of anticipated decision effectiveness. After each group reached a decision, its members individually responded to a short questionnaire asking for estimates of their decision's anticipated effectiveness (cf., London, 1977; Nutt, 1977). The questions were identical to those subsequently used by an independent team of judges and were internally consistent (coefficient alpha = .73). The average within-group interrater reliability was .71. The measure used was the average estimate of the five group members.

Group satisfaction with the decision process. Some researchers have used group satisfaction with the decision process as an index of decision effectiveness (Van de Ven & Delbecq, 1974). Although not synonymous with decision effectiveness, satisfaction with the process has been found to relate to member acceptance of a decision and its implementation (Bass, 1970). Group satisfaction with the decision process was measured by four questions answered on 7-point Likert scales (coefficient alpha = .80). The average within-group internal consistency was .77. The measure used was the average response to the satisfaction questions across the five participants.

Effectiveness rated by objective judges. Specific criteria to evaluate the effectiveness of each group's decision were established a priori by a panel of nine experts (see Stumpf et al., 1979, p. 775-776). Each had extensive education and experience in analyzing and coping with cases similar to the six decision problems in this study. Three of the six problems had actually been solved within the past eight years. Hence, objective data were available on the effectiveness of the decisions actually implemented.

After the specific criteria of an effective decision were established, an *independent group* of three raters used the criteria to evaluate the written decision of each experimental group. The three raters worked in a nominal mode. First, each rater responded individually to three questions addressing the overall effectiveness of the decision on a 7-point Likert scale. The raters then discussed the basis of their ratings; this was followed by a second individual rating. An overall rating was calculated by averaging the three raters' second ratings across all questions. The interrater reliability of the independent judges' rated effectiveness measure was .96.

Results: Study I

The results, averaged across six different decision situations, are shown in Table 1. A group's estimate of the anticipated effectiveness of its decision correlated strongly with its satisfaction with the decision process ($r = .83$, $p \leq .01$); but, *neither of these subjective proxy measures correlated with the*

hood of higher management approval were obtained in a post-decision interview. Each was asked: (1) to state how well he/she thought the candidate selected would succeed in the new position, and (2) to estimate the likelihood of higher management approval.

Objective, comparison measures. Criteria of effective promotion decisions were determined from a literature review and a field study of promotion practices (Stumpf & London, 1981). Positive candidate attributes included high past performance, high managerial potential ratings, favorable assessment center ratings, and recommendations. Two independent judges, using the promotion criteria, rated the effectiveness of each candidate selected. Interrater reliability was .70. The measure of rated effectiveness used was their average rating (coefficient alpha = .83).

Results: Study II

The means, standard deviations, and intercorrelations of the three criteria of decision effectiveness are shown in Table 2. The subjective, proxy measures correlated moderately ($r = .32$, $p \leq .05$). *The participant's estimates of anticipated effectiveness did not correlate significantly with the objective, comparison measure of effectiveness* ($r = -.14, .00$, ns). The mean for anticipated effectiveness was more lenient than rated effectiveness (4.6 versus 3.5) and had less variance (SD = .55 versus .83). (These measures cannot be statistically compared with *t*-tests and *F*-tests because the ratings are not made on identical scales.)

Table 2
Study II: Means, Standard Deviations, and Intercorrelations of Criterion Variables for a Management Promotion Decision Simulation

	Mean	SD	Anticipated Effectiveness	Likelihood of Decision Approval
Anticipated Effectiveness	4.6[a]	.55	--	
Likelihood of Decision Approval	2.2[b]	.83	.32*	--
Rated Effectiveness	3.5[a]	.83	-.14	.00

[a]These means are based on a 5-point scale, "5" being highest. $N = 72$.
[b]Likelihood of Decision Approval is based on a 3-point scale, "3" being highest.
*$p \leq .05$.

Method: Study III

Sample and Procedures

Three hundred and ninety-two middle- and upper-level managers participated in this study. Sixty groups were formed with five to eight members in each group. The study was a part of executive development programs, three to six weeks in length, offered by three different organizations: a major electronics company (32 groups), a major paper company (10 groups), and a northeastern graduate school of business (18 groups).

than their group's, 69 percent (192 of 279 managers) did worse; 31 percent (87/279) did better ($p \leq .001$). More than two-thirds of the participants inaccurately judged the quality of their individual revised decision relative to the quality of their group's decision. In this situation it would be difficult to make valid inferences regarding the objective accuracy from the individual manager's subjective estimate of the accuracy of his own decision.

Discussion

In response to our research question, there is a low correlation between subjective, proxy estimates and the objective, comparison measures of anticipated decision effectiveness. Further, the hypothesized leniency in participants' estimates was supported. The consistency of these results across studies is cause for concern. The rated effectiveness measures in Studies I and II approximated relevant, distal criteria and were more objective than the estimates of the subjects. These rated effectiveness measures were based on comparisons across groups (Study I) or individuals (Study II) confronting the same decision situation. Also, the more objective measures identified component facets of effectiveness and used specific criteria for each facet, an approach considered more valid than a single, global assessment of effectiveness (Smith, 1976).

Of equal concern is the low relationship between participant's estimates and objective measures in Study III that used actual data as the criterion. The low relationships suggest that making inferences about actual effectiveness from participant estimates may be misleading, if not erroneous.

Reasons for inaccurate participant estimates of anticipated decision effectiveness, while speculative, can be derived from the literature. For example, in Study I, groupthink may have had a pervasive effect on the participants (Janis, 1972). Having reached a difficult judgmental decision, the participants may have become cohesive and careless in their critical thinking. The participants in Study II probably did not develop a systematic model of the attributes and attribute weights for each candidate being considered for promotion. Information regarding candidates may have been overlooked or inappropriately evaluated. The participants in Study III who experienced differences and conflict may have become defensive and resistant. When left alone to make the decision, they may have distorted what they heard and reimposed their biases. It may also be that it is the implementation of a decision rather than the decision itself that is most important (Bass, 1970; Zand, 1981). Much of management consists of implementing decisions, which may mean adjusting, revising, and salvaging deficient decisions as their weaknesses emerge.

Limitations

Some possible research limitations should be noted. In Study I, group members could be confusing their estimate of the decision's anticipated effectiveness with their satisfaction with the decision process. The in-basket exercise in Study II included several acceptable candidates and managers

regarding aspects of the decision such as the relevant criteria, relative weights of the criteria, and the attributes of various decision alternatives to derive a preferred alternative based on a particular mathematical model (e.g., a subjective-expected utility model).

References

Bass, B.M. When planning for others. *Journal of Applied Behavioral Science,* 1970, *6,* 151-171.

Festinger, L., & Aronson, E. Arousal and reduction of dissonance in social context. In D. Cartwright and A. Zander, *Group dynamics: Research and theory.* N.Y.: Harper & Row, 1968.

Festinger, L., & Walster, E. Post-decision regret and decision reversal. In L. Festinger, *Conflict, decision, and dissonance.* Stanford, Calif.: Stanford University Press, 1964.

Hogarth, R.M., & Makridakis, S. Forecasting and planning: An evaluation. *Management Science,* 1981, *27,* 115-138.

Janis, I.L. *Victims of groupthink.* Boston: Houghton-Mifflin, 1972.

Lichtenstein, S., & Fischhoff, B. Do those who know more also know more about how much they know? *Organizational Behavioral and Human Performance,* 1977, *20,* 159-175.

London, M. Effects of information heterogeneity and representational roles on group member behavior and perceptions. *Journal of Applied Psychology,* 1977, *62,* 76-80.

London, M., & Stumpf, S.A. *Effects of candidate characteristics on management promotion decisions: An experimental study.* Technical Report, Basic Human Resources Research, American Telephone and Telegraph Company, Morristown, New Jersey, 1980.

Nutt, P.C. The merits of using experts or consumers as members of planning groups: A field experiment in health planning. *Academy of Management Journal,* 1976, *19,* 378-394.

Pfeiffer, J.W., & Jones, J.E. Consensus-seeking: A group ranking task. In *A handbook of structured experiences for human relations training* (Vol. II). Iowa City: University Associates Press, 1970, 22-25.

Schwab, D.P., Heneman, H.G., & DeCotiis, T.A. Behaviorally anchored rating scales: A review of the literature: *Personnel Psychology,* 1975, *28,* 549-562.

Smart, C., & Vertinsky, L. Designs for crisis decision units. *Administrative Science Quarterly,* 1977, *22,* 640-657.

Smith, P.C. Behaviors, results and organizational effectiveness: The problem of criteria. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology.* Chicago: Rand McNally, 1976, pp. 745-776.

Spool, M.D. Training programs for observers of behavior. A review. *Personnel Psychology,* 1978, *31,* 853-888.

Stumpf, S.A., Freedman, R.D., & Zand, D.E. Judgmental decisions: A study of interactions among group membership, group functioning, and the decision situation. *Academy of Management Journal,* 1979, *22,* 765-782.

Stumpf, S.A., & London, M. Management promotions: Individual and organizational factors affecting the decision process. *Academy of Management Review,* 1981, 6, 539-550.

Stumpf, S.A., Zand, D.E., & Freedman, R.D. Designing groups for judgmental decisions. *Academy of Management Review,* 1979, *4,* 589-600.

Thornton, G.C., III. Psychometric properties of self-appraisals of job performance. *Personnel Psychology,* 1980, *33,* 263-272.

Van de Ven, A.H., & Delbecq, A.L. The effectiveness of nominal, delphi, and interacting group decision making processes. *Academy of Management Journal,* 1974, *17,* 605-621.

Zand, D.E. Collateral organization: A new change strategy. *Journal of Applied Behavior Science,* 1974, *10,* 63-89.

Zand, D.E. *Information, organization and power.* New York: McGraw-Hill, 1981.

Received July 7, 1981■