A GROUP-BASED PROCEDURE FOR REVEALING JUDGMENTAL HEURISTICS AND BIASES

Ariel S. Levi, Indiana University at South Bend
Larry E. Mainstone, Valparaiso University

ABSTRACT

This paper describes a group-based exercise for helping participants: (1) understand their judgmental biases and the heuristics that produce them; (2) recognize overconfidence in their own subjective judgments, and why this overconfidence occurs; and (3) recognize the conditions which cause groups to improve (or worsen) judgmental accuracy.

INTRODUCTION AND PURPOSE

In recent years, researchers have uncovered a wide variety of deficiencies and biases in human judgment. Evidence has accumulated that people often depart systematically from optimal models of judgment and decision making. When estimating probabilities, judging the relationship between two variables, or assessing the utility of potential outcomes, individuals either fail to consider or use relevant information, or rely on information that is irrelevant from the standpoint of normative statistical models [1; 10; 15]

Such judgmental deficiencies and biases are largely due to individuals' reliance on heuristics, or mental strategies for processing information. Heuristics are generally efficient, and often produce reasonably accurate judgments and decisions [17]. However, in a substantial proportion of cases, they lead to systematically biased results.

The importance of judgmental bias is magnified by two findings. First, suboptimal performance in judgment and decision tasks is not confined to the poorly educated or to those lacking in expertise. The same biases and deficiencies have been demonstrated in experts and experienced professionals [8;18]. Second, people tend to be overconfident in the accuracy of their judgments [2;3;7]. This tendency reduces the likelihood that decision makers will critically examine a chosen course of action or the assumptions underlying it. In addition, overconfidence may hinder acceptance of techniques for improving judgment or decision makings

Given the potential costs of judgmental bias, it is not surprising that researchers and consultants have begun to develop and test methods for improving judgment [4;6;14;19]. A wide variety of judgment and decision aids are available. Most of these aids, which focus on the individual as the decision making unit, are based on the assumption that judgments and decisions can be improved if the individual learns how to structure the problem systematically according to some optimal statistical or decision model.

However, as Fischhoff and Johnson [6] and Fischhoff and Gotein [5] have pointed out, there are limits to the effectiveness of most current judgment and decision aids. These limits arise from several considerations. First, simply instructing individuals in theoretical knowledge of statistics or decision theory is insufficient for improving judgment. For improvement to occur, Individuals must learn when and how to apply the appropriate model, and must then become familiar with the subtleties of the model as it is applied in specific cases. Second, judgment and decision aids are often unwittingly "reactive," in that they may actually create or modify values and beliefs rather than simply reveal them. This reactiveness could modify the original problem in unintended yet consequential ways. Finally, judgment and decision aids do not necessarily provide the user with unambiguous answers. Frequently, a large dose of subjective judgment is necessary in order to apply the aid or interpret its results. Because individuals are not very accurate in assessing the extent of their own knowledge, they may be tempted either to rely too heavily on the aid, or to distort its results to accord more closely with their subjective judgment.

Judgment and decision aids will be effective to the extent that the user understands these limitations and takes them into account. Unfortunately, most individual-based decision aids do not deal with these obstacles to effectiveness. However, group-based judgment and decision aids may be more effective in helping individuals recognize and deal with these obstacles, a possibility that has begun to receive increasing attentions Nisbett & Ross [15], for example, suggested that open group discussion might reduce judgmental error by providing multiple perspectives on problems, and checks and balances against extreme individual judgments. Similarly, Shaughnessy [16] raised the possibility that group judgments and decisions might be less subject to the probabilistic misconceptions that affect individual judgments. If so, judgment and decision making might be taught most effectively in small groups.

Recent research on group processing of judgmental problems suggests that groups can be effective in revealing the heuristics that individuals use to make judgments. Open group discussion forces individuals to make explicit their assumptions and to specify the procedures they are using to arrive at their judgments. Group-based procedures thus do more than merely instruct individuals in theoretical knowledge; they require active analysis and the application of such knowledge.

Individuals can benefit from group processing of judgments, even though such processing does not always result in improved judgments (the evidence on this point is mixed; see [12; 13]). The confrontation between group members with differing viewpoints leads to more extensive processing of the judgment problem, and an appreciation of the potential validity of other viewpoints. In addition, compared to individual-based methods, active participation in group discussion often leads to more effective learning and retention [9; 11]. Thus groups have the potential to help individuals understand judgmental heuristics and biases, and recognize the importance of subjective judgment and unwarranted confidence in judgmental accuracy.

In this paper, we describe a group-based procedure for revealing judgmental heuristics and biases. We have found the procedure to be very effective for helping managers and management students gain an understanding of human judgment and its deficiencies. The specific purposes of the procedure are to: 1) enhance participants' understanding of judgmental biases and the heuristics that produce them; 2) help participants recognize overconfidence in their own subjective judgment, and to understand why this overconfidence occurs; and 3) help individuals recognize the conditions under which groups improve versus worsen judgmental accuracy.

THE PROCEDURE

The procedure includes four steps.

Step 1

In this step, participants are presented with a set of judgmental problems designed to illustrate the major heuristics and biases (the Appendix contains two sample problems). Each participant receives a packet containing about 8-10 problems to be done individually. Participants are to answer the problems using their best judgment, unaided by other individuals or sources of information. For each problem they are asked to write a brief description of how they arrived at their answers. In addition, for each problem, participants rate their degree of confidence in the correctness of their judgments on a 7-point bipolar scale.

Step 2

Participants meet in groups composed of 5-7 individuals. Each group discusses the problems and arrives an answer for each problem. The objective is for the groups to make better (i.e., more optimal) judgments than did the individuals. In processing the problems, there is no specific procedure the groups must follow; the procedure they use is up to the group members.

After the group has answered each problem, group members privately rate their individual confidence in the correctness of the group answer on a 7-point bipolar scale. In addition, for each problem, group members check-mark, in multiple-choice format, the process that the group used to make its judgment. There are four choices: consensus, majority vote, dominant individual(s), and other.

The length of time the groups meet depends on the number of problems in the problem set. A set of 8-10 problems will generally require about 45 minutes to an hours

Videotaping the groups as they work through the problems can provide some very interesting and useful process material for later discussion and analysis. It is possible to set up a stationary video camera, with a wide-angle lens, to tape the entire group. Participants adapt rapidly to the camera's presence; we have found that videotaping does not noticeably affect the group process.

At the end of the group meeting period, the instructor collects the individual judgments, the group judgments, and the confidence and group-process ratings, all of which have been recorded on appropriate forms.

Step 3

In this step, feedback is provided to the participants concerning their individual judgments, group judgments,

and confidence ratings. Feedback should be given within a week of the group meetings, so that participants will recall the problems and retain interest in them.

In order to provide effective feedback, it is necessary for the instructor to perform statistical analyses on the data. For each problem, these analyses should include:

1) means and modes of all the individual judgments;
2) means and modes of all the group judgments;
3) mean confidence ratings for the individual judgments;
4) mean confidence ratings for the group judgments;
5) mean differences between individual-judgment confidence ratings and group-judgment confidence ratings.

For more detailed feedback, the following analyses can be performed: comparisons between confidence ratings for correct versus incorrect individual judgments; comparisons between confidence ratings for correct versus incorrect group judgments; and examination of changes in confidence ratings within groups giving correct versus incorrect judgments.

In providing feedback, it is most useful to begin by posting summary data for each problem, organized in three columns: mean or mode of individual judgments; mean or mode of group judgments; and the correct judgment. Table 1 illustrates how the results of the problems in the Appendix would be presented.

TABLE 1
ILLUSTRATIVE SUMMARY RESULTS
FOR TWO JUDGMENTAL PROBLEMS

Problem	Individual judgment (n=124)	Group judgment (n=23)	Correct judgment
Hospital			
larger	18.5%	0%	
smaller	47.6%	82.6%	smaller
same	33.9%	17.4%	
% correct	47.6%	82.6%	
Cab			
mean	62.9	68.8	
mode	80.0	80.0	
			41.0
% correct	0	0	

Results presented in this manner enable participants to compare individual with group judgments. Here, it is apparent that groups were more accurate than individuals in the hospital problems In the cab problem, on the other hand, groups were no more accurate than were individuals.

Such results can serve as a starting point for discussing the following issues; Are group judgments generally more accurate than individual judgments?; Can groups worsen judgmental accuracy?; What determines whether groups will improve versus worsen accuracy?

Summaries of the confidence ratings should also be a part of the feedback. Mean confidence ratings should be presented for each problem. Two comparisons are particularly enlightening: confidence in individual versus group judgments; and confidence in correct versus incorrect judgments.

In general, confidence is higher for group than for individual judgments. Discussing judgmental problems in a group setting tends to increase confidence. This difference is usually quite slight, except when participants are very certain that their group has made the correct judgment. In the hospital problem, for example, average confidence in the group judgment was substantially higher than average confidence in the individual judgment, because several participants clearly recognized the relevance of a fundamental statistical principle, the law of large numbers. These individuals expressed their certainty to other group members, thereby increasing the general confidence level.

Confidence ratings often are not correlated with accuracy. In fact, confidence is usually about equally high for both correct and incorrect judgments. This finding is especially useful for leading participants to recognize the pervasiveness of overconfidence. Discussion can revolve around the following issues:
How do individuals assess the extent of their own knowledge, or the accuracy of their subjective judgment?; Why are individuals usually overconfident?; and, How can the tendency toward overconfidence be reduced?

Step 4

This step involves more extensive discussion and analysis of the judgmental problems, and the heuristics and biases that were revealed in the group meetings. In leading the discussion, it is helpful for the instructor to first list the heuristics that individuals used in making their judgments. These heuristics can usually be grouped on the basis of similarity into a small number of categories. Each category can then be analyzed and discussed in terms of how it compares to normative models of statistics or decision making.

Transcripts of group meetings are especially useful for helping participants understand different heuristics and the biases they produce. By examining the transcripts, participants can understand how different individuals, given the same information, can arrive at radically different judgments. Following is a partial transcript of a group discussing the cab problem.

B: Eighty percent.
A: You say 80%?
B: Yeah, because they tested the witness, and they came up with that probability of him being correct.
C: Well, I took it as the probability of 80 given 15, 80 I think it's lower than 80%.
D: I put 25%.
A: So did I, but I'm not sure why.
D: One thing I looked at, the witness' testimony, you'd have to throw that out in court, being incorrect 2 out of 10 times, no one's going to believe you.
B: That might be true in court, but here we're just trying to find the probability that the cab is blue.
E: My estimate was real low, I put 12%. I tried to use statistics...
F: I misread the question earlier. Now I agree with you [B], that it's 80%, because that's what the witness tests out to be.
C: I would have to disagree, because if you work it out mathematically, it would have to be lower, because you have to work in that only 15% of the cabs are blue.

B: What about the witness--are you just throwing out the fact that he said it was a blue cab?
C: That's the problem here--we're working with mathematical models, but we're also working with subjectivity, and if we can decide what mathematical model to go with, we could shade it according to the subjectivity.
A: I see, there's more green than blue, so it doesn't matter what the witness says, you have to count...
D: It matters how much you want to weigh the witness' statement.
B: You're running into two angles--he's right 80% of the time but 85% of the cabs are green. So you have to weigh one fact against the other.

This transcript illustrates how different perspectives on the problem emerged during group discussion. In this particular group, there were two "factions," one supporting a low estimate *(12-25%),* the other supporting a high estimate (80%). The confrontation between these two factions eventually led the participants to realize that both the witness' accuracy (80%) and the base rate of blue cabs (15%) were relevant items of information. Eventually, the group used an intuitive version of Bayes' theorem to arrive at an estimate of 46%, close to the correct estimate of 41%.

Participants enjoy going over the transcripts and comparing other groups' discussions with their own. By close analysis of the transcripts, participants can understand how different individuals processed the information to make their judgments.

APPENDIX IX
TWO SAMPLE JUDGMENT PROBLEMS

1. A certain town is served by two hospitals. In the larger hospital, about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower.

For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys. Which hospital do you think recorded more such days?

a) larger   b) smaller   c) about the same

2. A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

a) 85% of the cabs in the city are Green and 15% are Blue.

b) A witness identified the cab as a Blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs (half of which were Blue and half of which were Green), the witness made correct identifications in 80% of the cases and erred in 20% of the cases.

Question: What is the probability that the cab involved in the accident was Blue rather than Green?

Developments in Business Simulation & Experiential Exercises, Volume 13, 1986

REFERENCES

[1] Abelson, R. P., and Levi, A., Decision making and decision theory. In G. Lindzey and E. Aronson (Eds.), Handbook of social psychology (3rd ed.) (New York: Random House, 1985).

[2] Alpert, M. and Raiffa, H., A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, and A. Tversky (Ed.), Judgment under uncertainty: Heuristics and biases (pp. 294-305) (Cambridge: Cambridge University Press, 1982).

[3] Einhorn, H. J. and Hogarth, R. H. "Confidence in judgment: Persistence of the illusion of validity," Psychological Review, 1978, 85, pp. 349-4 16.

[4] Fischhoff, B., Debiasing. In D. Kahneman, P. Slovic, and A. Tversky (Eds.) Judgment under uncertainty: Heuristics and biases, pp. 422- 444 (Cambridge: Cambridge University Press, 1982).

[5] Fischhoff, B. and Gotein, B., "The informal use of formal models," Academy of Management Review, 1984, 9, pp. 505-512.

[6] Fischhoff, B. and Johnson, S., "The possibility of distributed decision making," presented at the Workshop on Political-Military Decision Making, The Hoover Institution, Stanford University, March 7-9, 1985.

[7] Fischhoff, B., Slovic, P.., and Lichtenstein, S.., "Knowing with certainty: The appropriateness of extreme confidence, Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, pp 552-564.

[8] Fischhoff, B., Slovic, P., and Lichtenstein) S., "Fault trees: Sensitivity of estimated failure probabilities to problem representation," Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, pp. 330-344.

[9] Fulmer, R. "Teaching the future: Using uninformed experts," presented at the 12th Annual Organizational Behavior Teaching Conference, Charlottesville, June 11-14, 1985.

[10] Kahneman, D., Slovic, P., and Tversky, A., Judgment under uncertainty: Heuristics and biases, (Cambridge: Cambridge University Press, 1982).

[11] Latham, G., and Wexley, K., Increasing productivity through performance appraisal, (Reading, MA: Addison Wesley, 1g81).

[12] Nagao, D. H., Tindale, R. S., Hinsz, V. B., and Davis, J. B., "Individual and group biases in information processing," presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, August 23-27, 1985.

[13] Neale, M. A., Bazerman, M. H., and Alperson, C. A., "Framing and the risky shift phenomenon: Yet another look)" presented at the 45th Annual Meeting of the Academy of Management, San Diego, August 11-14, 1985.

[14] Nisbett, R. E., Krantz, D. H., Jepson, C., and Kunda, Z., "The use of statistical heuristics in everyday inductive reasoning," Psychological Review, 1983, 90, pp. 339-363.

[15] Nisbett, R. E., and Ross, L.., Human inference: strategies and shortcomings of social judgment, (Englewood Cliffs, NJ: Prentice-Ball, 1980).

[16] Schaughnessy, J. M, The psychology of inference and the teaching of probability and statistics: Two sides of the same coin? In R. W. Scholz (Ed.) Decision Making under uncertainty) (Elsevier Science Publishers, North-Holland, 1983).

[17] Thorngate, W., "Efficient decision heuristics," Behavioral Science, 1980, 25, pp. 219-225.

[18] Tversky, A., and Kahneman, D., "Belief in the law of small numbers," Psychological Bulletin, 1971, 2, pp. 105-110.

[19] Wheeler, D. D., and Janis, I. L., A practical guide for making decisions, (New York: The Free Press, 1980).