Small-group polarization of judgments

I. R. ANDREWS

Department of Economics and Commerce
Simon Fraser University, Burnaby 2, B.C., Canada
and

D. L. JOHNSON

Bowling Green State University, Bowling Green, Ohio 43402

The polarizing effect of group discussion was tested with single-cue and multiple-cue risk-free stimuli. As predicted, ratings by groups were more extreme than ratings by individuals when judgments were based upon favorable or unfavorable information cues, but not when ratings were based upon neutral information cues. It was suggested that if group judgments are more extreme than individual judgments on both risk-free and risk-salient stimuli, then it might be advantageous to treat the "risky shift phenomenon" as a special subcase of a more general phenomenon.

Explanations of the risky shift phenomenon—the tendency for individuals to choose higher risk alternatives after discussion in small groups—have assumed that risk is the essential feature of the complex stimulus situation. For example, Brown (1965) proposed that the shift occurs because risk is more of a cultural value in North America than is conservatism. Kogan & Wallach (1967) have argued that it is the diffusion of responsibility for the risky decision which leads to the shift.

Unfortunately, the power of these risk-emphasizing explanations has been reduced by two recent streams of evidence. First, there are a number of studies in which group discussion led to a conservative shift in risk choices (e.g., Rabow et al, 1966; Stoner, 1968; Johnson & Andrews, 1971), an outcome not readily explained in terms of "risk as a cultural value" or "diffusion or responsibility for a risky choice." Second, judgmental shifts after group discussion have been produced with attitude statements containing no element of risk (Moscovici & Zavalloni, 1969; Doise, 1969; Myers & Bishop, 1970). This second stream of evidence suggests the possibility that the risky shift is a special subcase of a more general phenomenon, namely, the tendency for small group discussion to polarize judgments about many kinds of stimuli.

In the present study, the tendency of small groups to form more extreme judgments than do individuals was tested with both single- and multiple-cue stimuli. The cues provided were favorable, neutral, or unfavorable descriptive statements taken from a list standardized by Uhrbrock (1961). These statements contain no element of risk.

It was predicted that small-group ratings would be lower than individual ratings when they were based on

unfavorable descriptive statements and higher than individual ratings when they were based on favorable descriptive statements. For ratings based on neutral descriptive statements, no difference between small groups and individuals was expected, since there would be no initial trend to be polarized by small-group discussion

SUBJECTS AND PROCEDURES

A total of 96 male and 58 female undergraduates were run in same-sex groups of 3, 4, or 5 each—or as individuals. No S participated under both conditions. The individual Ss were run 3, 4, or 5 at a time in the same laboratory where small groups were run. Assignment to groups and to conditions was random, with the exception that groups included either all male or all female Ss. The final data revealed no difference due to sex or group size. For data analysis, individuals who were run at the same time were treated as concocted groups to maintain equal entries in each cell.

Since the task required Ss to rate university teachers on the basis of the descriptive statements provided by the E, relevance to teaching performance was the key criterion for selecting

statements from Uhrbrock's (1961) standardized list. The final selection provided eight favorable, eight neutral, and eight unfavorable statements to serve as single-information cues. Multiple-cue stimuli were produced by putting four descriptive statements on the same information card. Of primary interest to this study were the homogeneous cards, four with all favorable statements, four with all neutral statements, and four with all unfavorable statements. Twelve nonhomogeneous multiple-cue cards were included as fillers.

The Ss were asked to use these single-information and multiple-information cues to rate university teachers on an 11-point scale anchored with statements such as "extremely good teacher," "average teacher," etc. Individuals made the ratings individually, but in the presence of others, whereas groups were asked for a group consensus decision. On the multiple-cue stimuli, it was specified in the initial instructions that "...all four descriptions are accurate and each is equally important."

RESULTS

As predicted, for the single-cue judgments (Table 1), ratings by groups were lower than ratings by individuals on 7 out of 8 unfavorable statements and higher than ratings by individuals on 6 out of 8 favorable statements. Results for the homogeneous multiple-cue stimuli (Table 2) were even more consistent with the predictions—ratings by small groups were lower than ratings by individuals on 4 out of 4 unfavorable cue sets and higher than ratings by individuals on 4 out of 4 favorable cue sets. Combining single- and multiple-cue results for both favorable and unfavorable descriptive statements, ratings by small groups were more extreme than were ratings by individuals on 21 out of 24 stimulus cards.

In the above analysis, the data were collapsed across S samples to produce

Table 1
Ratings of Teacher Performance Based Upon Single Cue Descriptive Statements

Favorable Cue Number	Ratings by Individuals (N = 18)*	Ratings by Groups (N = 18)	Unfavorable Cue Number	Ratings by Individuals (N = 18)*	Ratings by Groups (N = 18)
1	8.60	9.67	1	2.36	1.58
2	9.76	10.25	2	1.81	1.00
3	8.70	8.92	3	2.61	1.58
4	9.18	9.75	4	1.65	1.08
5	8.86	8.17	5	2.49	2.08
6	7.82	6.83	6	2.47	1.25
7	9.04	9.25	7	2.40	2.67
8	9.05	9.50	8	1.47	1.25
Mean	8.88	9.05	Mean	2.16	1.63
Variance†	0.41	0.73	Variance†	0.25	0.14

*Individuals were treated as 18 concocted groups for statistical analysis.

†Variance among the mean ratings by 18 concocted groups or by 18 actual groups.

Table 2
Ratings of Teacher Performance Based Upon Multiple Cue Descriptive Statements

Favorable Cue Number	Ratings by Individuals (N = 18)	Ratings by Groups (N = 18)	Unfavorable Cue Number	Ratings by Individuals (N = 18)	Ratings by Groups (N = 18)
1	9.47	10.00	1	1.40	1.00
2	9.89	10.54	2	1.44	1.00
3	9.43	10.50	3	1.22	1.00
4	9.84	10.46	4	1.48	1.04
Mean	9.66	10.37	Mean	1.39	1.01
Variance	0.30	0.90	Variance	0.14	0.01

individual and group mean ratings for each stimulus card. In a second analysis, the data were collapsed across the stimuli within an experimental cell to produce a separate mean rating for each of the 18 concocted groups (individuals) and for each of the 18 actual groups. Variances among these mean ratings and the associated overall means are shown at the bottom of Tables 1 and 2. For the single- and multiple-cue data, a split-plot ANOVA revealed that the two main effects and the interaction term were significant at the .05 level. Since the significant interaction terms confirmed that the group vs individual main effect was different at the three-statement category levels, it was possible to proceed with a separate preplanned t test for each category level of the single- and multiple-cue statements.

The mean ratings by groups and individuals were significantly different on the single-cue unfavorable statements ($t = 2.45$, $p < .05$) and on the multiple-cue favorable statements ($t = 2.65$, $p < .05$). Differences between the mean ratings by groups and individuals were in the directions predicted for the single-cue favorable statements and for the multiple-cue unfavorable statements, but the t values were not significant at the .05 level. The authors believe that the difference between groups and individuals would have been significant on the multiple-cue unfavorable statements if the scale had allowed groups to assign ratings below 1.00. This belief is based on the fact that all but 1 of the 18 actual groups had a mean rating on these stimuli which was lower than the lowest mean rating among the 18 concocted groups. On balance, it seems reasonable to conclude from the above analyses that the tendency for groups to make more

extreme decisions than individuals was not simply a function of random chance.

Instead of clustering as expected around the center of the 11-point rating scale, individual responses to the neutral descriptive statements produced mean ratings slightly below the scale center. This slight initial trend toward the negative side was polarized by small-group discussion to the point where the mean difference between individual and group ratings was significant at the .05 level. Consistent with our hypothesis, however, this polarization effect disappeared with the four multiple-cue stimuli. In fact, the mean ratings for groups and individuals across the four neutral-level multiple-cue stimuli were almost identical. Thus, the data lend some support for the notion that the polarizing effect of group discussion is less apt to occur on stimuli which do not evoke a clear trend away from the center of the judgment scale.

DISCUSSION

Previous research has demonstrated that judgments by individuals become more extreme when these same individuals are influenced by group discussion. The present study extended that conclusion to a between-Ss design where different Ss were used for the individual and group conditions. More importantly, it provided additional support for the idea that the polarizing effects of group discussion occur for many kinds of stimuli and not just on stimuli which contain an element of risk. As suggested earlier, if group judgments are more extreme than individual judgments on both risk-free and risk-salient stimuli, then it may be advantageous to treat the risky shift as a special subcase of a more general phenomenon.

Still unanswered, however, is the important question of why group discussion leads to more extreme judgments. Elsewhere the authors (Johnson & Andrews, 1971) have suggested that group discussion could clarify the characteristics of the stimuli being judged and might also lead to a better understanding of the role demands of the situation. In addition, there might be a climb-on-the-bandwagon effect when group members discover that most of them are leaning in the same direction on a particular question. Though not admissible as rigorous evidence, one of the authors has observed a pronounced bandwagon tendency in promotion and tenure committee meetings. Once a clear trend for a yes or no vote has been established, evidence contrary to that trend tends to be repressed, while evidence supporting the trend tends to be exaggerated. Lastly, there is the possibility that distortion in favor of a prevailing trend is furthered by errors in cue perception or in the cue combination process.

If future research confirms the polarization effect as a reliable phenomenon for a wide variety of stimuli and situations, there would be no shortage of explanatory mechanisms to be tested.

REFERENCES

- BROWN, R. *Social psychology*. New York: Free Press, 1965.
- DOISE, W. Intergroup relations and polarization of individual and collective judgments. *Journal of Personality & Social Psychology*, 1969, 12, 136-143.
- JOHNSON, D. L., & ANDREWS, I. R. The risky shift hypothesis tested with consumer products as stimuli. *Journal of Personality & Social Psychology*, in press.
- KOGAN, N., & WALLACH, M. A. Risk taking as a function of the situation, the person, and the group. In G. Mandler (Ed.), *New directions in psychology III*. New York: Holt, Rinehart, & Winston, 1967.
- MYERS, D. G., & BISHOP, G. D. Discussion effects on racial attitudes. *Science*, 1970, 169, 778-779.
- MOSCOVICI, S., & ZAVALLONI, M. The group as a polarizer of attitudes. *Journal of Personality & Social Psychology*, 1969, 12, 125-135.
- RABOW, J., et al. The role of social norms and leadership in risk taking. *Sociometry*, 1966, 29, 16-27.
- STONER, J. A. F. Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 1968, 4, 442-459.
- UHRBROCK, R. S. 2,000 scaled items. *Personnel Psychology*, 1961, 14, 375-420.