The Group Polarization Phenomenon

David G. Myers
Hope College

Helmut Lamm
Universität Mannheim, West Germany

Experiments exploring the effects of group discussion on attitudes, jury decisions, ethical decisions, judgments, person perceptions, negotiations, and risk taking (other than the choice-dilemmas task) are generally consistent with a "group polarization" hypothesis, derived from the risky-shift literature. Recent attempts to explain the phenomenon fall mostly into one of three theoretical approaches: (a) group decision rules, especially majority rule (which is contradicted by available data), (b) interpersonal comparisons (for which there is mixed support), and (c) informational influence (for which there is strong support). A conceptual scheme is presented which integrates the latter two viewpoints and suggests how attitudes develop in a social context.

In 1961, James Stoner discovered a reliable but nonobvious effect of group discussion. To examine the popular notion that groups are more cautious and less daring than individuals, Stoner had six persons at a time respond as individuals to a series of story problems called "choice-dilemmas" items (developed by Kogan & Wallach, 1964). The subject's task was to advise the fictional character of each item as to how much risk he should take in facing a given decision dilemma. A sample problem illustrates the task.

George, a competent chess player, is participating in a national chess tournament. In an early match he draws the top-favored player in the tournament as his opponent. George has been given a relatively low ranking in view of his performance in previous tournaments. During the course of his play with the top-favored man, George notes the possibility of a deceptive though risky maneuver which might bring him a quick victory. At the same time, if the attempted maneuver should fail George would be left in an exposed position and defeat would almost certainly follow.

Imagine that you are advising George. Please check the lowest probability that you would consider acceptable for the risky play in question to be attempted.

George should attempt to play if the chances are at least

<table>
<thead>
<tr>
<th>Probability</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>1 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>2 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>3 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>4 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>5 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>6 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>7 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>8 in 10</em></td>
<td>play</td>
</tr>
<tr>
<td><em>9 in 10</em></td>
<td>play</td>
</tr>
</tbody>
</table>

George should attempt the play only if it is certain (i.e., _10 in 10_) that the play would succeed.

After individually marking their advice on all the items, the participants then assembled as a group and discussed each item until they agreed. The finding that groups were on the whole more risky than their average individual member (generally by about one scale unit or less) was immediately dubbed the risky-shift phenomenon and was followed by a wave of investigations of group risk taking. These studies have taken place in a dozen different nations, indicating that Stoner's results are not peculiar to his subject population. They also reveal that group decision making is not an essential component of the procedure; a brief period of discussion followed by individual responses will also produce a shift in the group average.
This mass of risky-shift studies conducted during the 1960s, and thought by some (e.g., Smith, 1972) to have been a fruitless fad, is now appearing to have had more heuristic value than previously realized. This paper summarizes recent experiments examining the generality of the group shift effect uncovered in the risky-shift experiments, and it reviews recent efforts to construct a theoretical home for the group shift phenomenon.

The risky-shift literature, which is foundational to this review, has been well summarized elsewhere (Cartwright, 1971; Clark, 1971; Dion, Baron, & Miller, 1970; Pruitt, 1971a, 1971b; Vinokur, 1971b). Hence, we do not again summarize data pertaining to the conditions under which group shift occurs on choice-dilemmas items or to theoretical approaches which have been discounted and are no longer actively pursued.

**ORIGIN AND DEFINITION OF THE GROUP POLARIZATION CONCEPT**

It is now widely recognized that the designation risky shift was a misnomer that unfortunately induced many investigators to perceive the phenomenon from the perspective of the dependent variable, risk taking, rather than to think in broader theoretical terms about the effects of intragroup communication on attitudes and behavior. The risky-shift label continued to guide experimentation long after it was well established that shift to greater caution could be reliably demonstrated on certain choice-dilemmas items. In fact, one powerful and important finding from the choice-dilemmas studies is that the individual items differ from one another in (a) mean initial response and (b) mean shift and that a and b are very highly correlated (generally .70 to .90 [e.g., Teger & Pruitt, 1967]). It is important to note that this correlation is occurring at the level of the item, with data averaged over groups. Items which elicit relatively risky initial tendencies generally elicit further shift toward the risky extreme after discussion. Items with relatively cautious initial means are more likely to elicit further shift in the cautious direction.

Restating this finding in more general terms we have the group polarization hypothesis: The average postgroup response will tend to be more extreme in the same direction as the average of the pregroup responses. The term group polarization originates from the writings of Serge Moscovici and his colleagues (e.g., Moscovici & Zavalloni, 1969). Note that polarization refers to an increase in the extremity of the average response of the subject population. (This use of polarization in the relevant literature is somewhat specialized, as in general usage the term may also refer to a split within a group of people.)

This polarization concept should also be distinguished from a related concept, extremization. Whereas polarization refers to shifts toward the already preferred pole, extremization has been used to refer to movement away from neutrality, regardless of direction. Since all instances of group polarization are instances of extremization, but not vice versa, extremization may be easier to demonstrate than polarization.

It should also be noted that conclusions about group polarization need not apply to individuals. As Fraser, Gouge, and Billig (1971) illustrated, a three-person group with scores of +3, +1, and −2 on some +3 to −3 attitude scale might all put down +1 after discussion. At the group level, a slight move to extremity (which in this case is also polarization) has occurred in the shift from a mean of +.67 to +1. But none of the individuals is more extreme than prior to the discussion. In fact, since group members generally converge, this pattern of responses is fairly typical.

Finally, it should be noted that defining group polarization as an enhancement of the tendency initially dominant in the whole sample requires the definition of a subjective neutral point on a bipolar scale. Very often the definition of a psychological midpoint (where there is no dominant tendency) is no simple matter. On the choice-dilemmas items, for examine, it probably is not 5 chances in 10, or even 5.5 in 10 (the mathematical midpoint of the scale). For example, post hoc regression analyses indicate that items with an initial mean between 6 and 7 in 10 will generally not yield much shift, while those with initial means below 6 will generally yield
risky shift and those with initial means above 7 will often elicit shift toward the higher probability pole (e.g., Myers & Aronson, 1972). Post hoc regression functions are, of course, a circular way of defining neutral points to then use in testing the group polarization hypothesis. More direct data collected by Miller (Note 1) indicate that with items on which subjects' positions are located near the mathematical midpoint of the scale (5.5), subjects nonetheless tend to verbalize a preference for the risky alternative when asked to advise whether the protagonist should take the risk.

We now examine group discussion experiments using measures other than the choice dilemmas to see whether the group polarization hypothesis derived from the choice-dilemmas literature has external validity with other response dimensions. The generalizability of the phenomenon has theoretical as well as empirical and practical relevance. If group polarization can be demonstrated with a variety of measures and be given a satisfactory general explanation, it will make unnecessary those explanations of the risky-shift phenomenon which are risk specific, as well as those which attribute the shift effect to peculiarities in the instructions or the metric of the choice-dilemmas items.

**GENERALITY OF GROUP-INDUCED POLARIZATION**

For convenience we have organized these studies into seven categories: attitudes, jury decisions, ethical decisions, judgments, person perceptions, negotiation behavior, and risk measures other than the choice dilemmas. This categorization is admittedly somewhat arbitrary. Within each category we briefly present two or three illustrative studies and then summarize findings from the remaining relevant studies.

**Attitude Studies**

*Varying the stimulus materials.* Two different research paradigms have yielded data relevant to the group polarization hypothesis. The first strategy has been to engage subjects in discussion of stimulus materials which have elicited some dominant predisposition and then to determine whether talking in groups tends to enhance the dominant leaning (just as choice-dilemmas items have elicited risky or cautious initial tendencies and corresponding shift tendencies).

Most studies of group discussion in social psychology and in speech-communications were designed for purposes other than examining a strengthening of the initial opinion average. Hence they report no prediscussion opinion average or they involved an initial opinion average near the neutral point (e.g., McCauley, 1972; Utterback, 1950). In other studies, however, it is possible to infer the direction of initial preferences. Robinson (1941) conducted lengthy discussions of two attitudes. On attitude toward war, where students were initially quite pacifistic, there was a nonsignificant shift to even more pacifism following discussion. On attitude toward capital punishment, to which students were initially opposed, there was a significant shift to even stronger opposition. Most other studies predating the contemporary group shift literature also evidence some degree of polarization effect (Geier, Forston, & Larson, 1970; Hopkins, 1964, p. 131; Utterback, 1954), but not Miller and Biggs (1958).

More contemporary attempts to extend group shift findings to the attitude realm have met with consistent success. Moscovici and Zavalloni (1969) observed that French students' initially positive attitudes toward DeGaulle and negative attitudes toward Americans were strengthened through discussion. Doise (1969b) reported that discussion enhanced the negative attitudes which French architectural students had toward their school. Judging from the number of citations these two papers have received, they appear to have been a stimulus for much of the research which followed. Subsequent studies investigating attitudes regarding social issues (Gouge & Fraser, 1972; Paicheler & Bouchet, 1973) and life situation dilemmas (Myers & Bishop, 1971) have yielded similar polarization effects.

*Group composition.* Another set of attitude experiments has explored group polarization using a strategy dissimilar to that of the risky-shift paradigm. Groups were composed of subjects sharing common inclinations rele-
vant to the discussion materials and their average shift was contrasted with that of groups of subjects sharing the opposite tendency. The group polarization prediction was that discussion with similar others will increase the attitude gap between separated homogeneous groups.

Myers and Bishop (1970) composed homogeneous groups of relatively high-, medium-, or low-prejudice subjects, using a racial attitude inventory. Group members then responded to eight new racial attitude items before and after discussion. The discussion with others having similar racial attitudes significantly increased the gap between the high- and low-prejudice groups. Mitnick and McGinnies (1958) observed a similar phenomenon using racial attitude materials.

Since there is ample evidence that people do in fact prefer being with others who have similar attitudes and values (Byrne, 1971), the question of whether separation on the basis of shared values generally produces increased intergroup polarization is of more than mere academic interest. Attempts to replicate the intergroup polarization effect have met with uneven success. Myers and Bach (1974) did not observe increased polarization between separated groups of pacifistic and militaristic subjects after discussion of items relevant to their value differences. Rather, all groups significantly increased in pacifism. However, Myers (1975) did observe that groups which were liberal in attitudes regarding the role of women polarized further from groups which were traditional after discussing relevant items. And Cvetkovich and Baumgardner (1973) found that groups of subjects who were initially nonpunitive in their attitudes toward civil disobedience became even more tolerant after discussion, while more punitive groups did not. In these studies, and some others to be considered later, the general trend of shift averaged over all conditions seems consistent with the prevailing external norm.

Field observations of social polarization appear consistent with the laboratory studies of intergroup polarization. Coleman (1957) concluded from his analyses of opinion polarization during community conflict that group discussion . . . is such an important phenomenon in community controversies that in the case studies examined most descriptions of behavior during the intense part of the controversy were descriptions of discussion and of attempts to persuade or reinforce opinion. (p. 18)

Homogeneous grouping was a source of community polarization and the occurrence of social conflict further heightened the proliferation of associations among those who feel one way, and the attenuation of association between those who feel differently. One's statements meet more and more with a positive response; one is more and more free to express the full intensity of his feeling. (p. 14)

Research on student change during college, summarized by Feldman and Newcomb (1969), reveals even clearer evidence of intergroup polarization.

Initial differences among students in different colleges and in different curricula are accentuated or amplified as students progress through college. Instances of this same phenomenon also occur with respect to initial differences among students entering different types of residences. (p. 209)

For example, the tendency for fraternity members to be more conservative and prejudiced than independents tended to be smallest at the freshman and sophomore levels and largest at the senior level. Feldman and Newcomb surmised that this may occur partly because "the reciprocal influences of members on one another reinforce and strengthen extant orientations (p. 223)." The likelihood of this type of communication effect is one of the reasons Pettigrew (1969) has argued so strongly that increased racial separation is antithetical to racial reconciliation.

To summarize, the weight of the evidence from attitude studies strongly confirms the group polarization phenomenon when stimulus materials are used which elicit a dominant initial tendency in the sample population, and the evidence is generally indicative of intergroup polarization effects when the group composition paradigm is used.

Jury Decisions

Juries are small groups which are given the task of arriving at a decision following discussion. Do decisions following jury delibera-
tion differ in any predictable way from the average of the predeliberation opinions of individual jury members? Three small sets of studies, the first two involving experimental simulations, are summarized.

**Varying the stimulus materials.** Myers and Kaplan (1976) engaged their subjects in discussion of stimulus materials which elicited a dominant predisposition of guilty or not guilty. After discussing traffic cases in which the defendants were made to appear as low in guilt, the subjects were even more definite in their judgments of innocence and more lenient in recommended punishment. After discussing “high-guilt” cases, the subjects polarized toward harsher judgments of guilt and punishment.

Izzett and Leginski (1974) observed, however, that an initial tendency for unattractive defendants to receive harsher sentences than did attractive defendants was reduced rather than exacerbated by discussion. Foss and Foss (Note 2) and Rumsey and Castore (Note 3) also observed greater leniency after discussion, although Heimbach (Note 4) observed greater harshness by groups. In all four of these experiments, however, it is difficult to describe the subjects as initially tending toward leniency or harshness, since no psychological neutral point can be defined.

**Group composition studies.** Vidmar (Note 5) composed groups of jurors high or low in dogmatism. The high-dogmatism juries shifted toward harsher sentences following discussion, and the low-dogmatism groups shifted toward more lenient sentences, despite the fact that the high- and low-dogmatism juries did not differ in their predeliberation judgments. Laughlin and Izzett (Note 6) observed that groups composed of subjects which were attitudinally similar to the defendant shifted toward greater leniency following group discussion, whereas subjects who were attitudinally dissimilar did not shift.

**Actual court decisions.** These findings, from experimental simulations of the jury process, do not lend themselves to clear-cut generalizations. They do, however, suggest the possibility that response-enhancement processes may operate in judicial group settings. Walker and Main (1973) provided more direct evidence. They compared civil liberties decisions by individual federal district court judges to civil liberties decisions by three-judge panels. The group condition produced substantially more libertarian decisions than did the single-judge condition (65% versus 30%). A subset of these decisions also involved rulings of the constitutionality of statutes. Main and Walker (1973) observed that these constitutionality decisions were also more libertarian in the group condition (65% versus 45%). Although a minority of the single-judge decisions were prolibertarian, Walker and Main surmised that the preexisting private values of the judges were actually prolibertarian and that their decisions made alone were compromised in the face of antilibertarian public pressure. Their private values were then supposedly released and reinforced in the professional group context.

Kalven and Zeisel (1966) presented evidence from 225 trials that is consistent with the group polarization hypothesis (the initial majority predicts the consensus outcome 90% of the time), and they concluded that the deliberation process might well be likened to what the developer does for an exposed film: it brings out the picture, but the outcome is predetermined. . . . From what we have been able to perceive thus far, the process is an interesting combination of rational persuasion, sheer social pressure, and the psychological mechanism by which individual perceptions undergo change when exposed to group discussion. (p. 489)

**Ethical Decisions**

Our third category of experiments deals with ethical and moral judgments. Horne and Long (1972) engaged their subjects in discussion of situations involving conflicts between universalistic and particularistic moral obligations. The subjects shifted toward increased universalistic following discussion, especially on items on which universalism was initially valued. Alker and Kogan (1968) asked their subjects to predict their own behavior in an ethical conflict situation. They observed an initial leaning away from the universalistic (ethical) pole and an enhancement of this tendency after discussion. Myers, Schreiber, and Viel (1974) used some simple ethical-legal dilemmas (e.g., a scenario in which a mature-looking but underage teenager is deciding whether to order a drink). The subjects,
male felons residing at a correctional institution, initially favored the illegal behavior and favored it even more following discussion.

Two other studies concern altruistic behavior. Schroeder (1973) had his subjects advise protagonists as to how much time or money they should sacrifice for a particular need. The altruistic direction, initially socially desired, was enhanced by group interaction. Baron, Roper, and Baron (1974) asked their subjects how much of their own money they would pledge to Bengali relief and how much they would recommend their student body contribute. Although generosity was presumed to be socially desirable (because the subjects perceived themselves as more generous than others), the pledges were actually more stingy after group discussion. (It might be noted, however, that actual prediscussion pledges—less than $2.00 per person—were also on the stingy side.) The idea of contributing from student body funds elicited greater generosity initially and a trend toward further increases after discussion. Both of these studies are of ambiguous relevance to the group polarization hypothesis, however, as it is impossible to define a 0 point on the altruism–selfishness continuum.

Finally, a series of studies by Rettig and his associates (Chapko, 1972; Rettig, 1966, 1972; Rettig & Turoff, 1967; Rettig, Note 7) has examined the effects of discussion on predictions of ethical risk taking. Rettig has suggested that his findings of increased predicted risk taking following group discussion run contrary to theorizing about the risky shift, since ethical risks are socially deviant. But close scrutiny of his data reveals that the discussions enhanced predicted ethical risk tendencies when and only when risk was of high reinforcement value (money stolen or "borrowed" was needed for a crucial medical operation). In this situation, the individuals were more likely to justify the illegal behavior, and social interaction further increased its perceived desirability.

**Judgment**

This set of studies centers more on judgments of fact than on social evaluations. As such, it borders on the group problem-solving literature, which we do not review. Kogan and Wallach (1966) compared the average absolute extremity of individual pregroup judgments (e.g., "What are the chances that an American family owns its own house?") with the extremity of consensus judgments. The average consensus was not more extreme than the mean absolute extremity of initial judgments. Doise (1971) reanalyzed these data in search of group polarization that may nonetheless have been occurring (enhancement of the initial algebraic mean of responses to an item). He found that "when the initial polarization is strong enough, the consensus significantly moves toward the pole that already attracted the individual responses (p. 515)."

Moscovici and Nève (1973) observed a similar result with social judgments.

Johnson and Andrews (1971) had their subjects judge their probable satisfaction with new consumer products that were preselected for high or low desirability. The desirable products were judged as even more satisfactory after discussion; the opposite was found for the undesirable products. However, Semin and Glendon (1973) intensively studied one real business committee and found that their judgments of the relative importance of different job criteria were unaffected by discussion. Their procedures made the pregroup judgments highly salient, however.

In judgments of the actual probability of success on choice-dilemmas items, Madaras and Bem (1968) and Lamm, Trommsdorff, and Kogan (1970) observed greater pessimism after group discussion. It is difficult to fit the group polarization hypothesis to their results, since the subjects were generally near the middle of the pessimism–optimism scale prior to discussion. Lamm and Trommsdorff (1974) failed to replicate the pessimistic shift with probability judgments concerning social and political change. In this study, too, the subjects were generally near the middle of the scale prior to discussion. Finally, Vidmar (1974) reported broader category-width judgments by groups than by individuals. Across all items, individuals already tended toward "broadness" in total score. But contrary to the polarization hypothesis, the shift to broader categories was not greatest on the individual items, on which initial judgments were broadest. Thus, while there is some evidence...
of group polarization in studies of factual judgment, the phenomenon is not as reliably confirmed as in studies requiring a greater degree of social evaluation.

**Person Perception**

Another set of experiments has explored the effects of group interaction on evaluations of hypothetical or real stimulus persons. Johnson and his associates engaged students in an evaluation of faculty. In one experiment (Andrews & Johnson, 1971), the subjects were provided with favorable or unfavorable cues describing a hypothetical faculty. With favorable cues, groups were more positive than individuals and with unfavorable cues, they were more negative. However, Shrewsberry and Johnson (Note 8) failed to replicate this effect. Subsequent studies (Johnson, Note 9; Krapf, Note 10) engaged students individually and then as groups in judgments of their actual professor. Groups tended to be more negative than individuals when the average individual ratings were already somewhat negative and more positive when individual ratings were positive. Myers (1975) also used a faculty evaluation task. The subjects responded to 200 word descriptions of "good" or "bad" faculty with a scale judgment and by distributing a pay increase budget among the hypothetical faculty. As predicted by the group polarization hypothesis, good faculty were rated and paid even more favorably after the group interaction, and contrariwise for the bad faculty.

Moscovici, Zavalloni, and Louis-Guerin (1972) and Moscovici, Zavalloni, and Weinerberger (1972) reported increased extremization following group discussion of impression formation materials (descriptive adjectives and photographs). However, they presented no data enabling one to determine whether this was also a polarization effect. Doise (1970) used paragraphs portraying a stimulus person as relatively extroverted or introverted and did observe greater polarity in group impressions of extroversion-introversion. Thus, in general, the work on person perception supports the group polarization hypothesis, especially when the stimulus materials are more complex than just a single adjective.

**Negotiation and Conflict**

Might group interaction effects also extend to bargaining and negotiation situations? The available data are minimal but intriguing. In a labor-management simulation, Rabbie and Visser (1972) instructed union bargaining teams to set expectation levels individually and then as groups. Individuals set higher aspiration levels when the issue was important and when the bargaining position was strong and lower aspiration levels when the issue was less important and when the bargaining position was weak. These tendencies were strengthened in the group condition.

Lamm and Sauer (1974) observed a similar polarization in a bargaining experiment which involved real consequences. When asked to distribute 18 profit units between themselves and another player, individuals initially proposed giving 64% to themselves. After discussing their positions with other individuals (whom they were not competing against), they increased their demand to almost 70%. Control subjects retested after private study showed no change.

Myers and Bach (1976) compared the conflict behavior of individuals and groups, using an expanded prisoner's dilemma matrix cast in the language of a gas war. There was no difference in their conflict behavior (both individuals and groups were highly non-cooperative). But on postexperimental scales assessing the subjects' evaluations of themselves and their opponents, individuals tended to justify their own behavior, and groups were even more inclined toward self-justification. This demonstration of group polarization supports Janis's (1972) contention that in situations of intergroup conflict, group members are likely to develop a strengthened belief in the inherent morality of their actions.

**Risk Taking**

The general finding of risky shift on choice-dilemmas items stimulated numerous experiments in which group-induced risk taking was sought on other risk measures. Since it was risky shift and not group polarization that was being sought, many of these experiments used betting materials in which the expected value of all outcomes was near zero.
and the initial mean did not suggest a dominant prediscussion trend toward either risk or caution. Not surprisingly, a risky shift was often not realized in these studies, causing investigators to question the generality of risky shift. Burnstein (1969) and Doise (1969a) reviewed research on skill and chance risk measures and observed, consistent with the group polarization hypothesis, that in those experiments which did produce a risky shift, initial responses tended to be risky.

More recent research can also be seen to fit the polarization process. Manipulating the expected value of bets such that the taking of increased risk would increase the expected value of one's outcome has substantial effects on group-induced shift (Davis, Kerr, Sussman, & Rissman, 1974; Marquis & Reitz, 1969). When the expected value of the bet is positive, groups will take more risk than will individuals. When the expected value is negative, groups will take less risk than will individuals.

When the expected value of the bet is held at zero, then the direction of the shift may be seen to vary with other parameters in accord with the affect of these parameters on pregroup betting tendencies. Using a procedure adapted from Pruitt and Teger (1969), Zaleska (1974, 1976) and Lamm and Ochsman (1972) we observed that group risk taking exceeded that by individuals only when the stake (potential loss) was small. Zaleska found that it is in precisely this condition that individual pregroup inclinations also tend to be quite risky. When the stake was large (and the potential gain also high), individuals were more cautious and groups exceeded them in caution.

A reanalysis of the Lamm-Ochsman data (Lamm, Myers, & Ochsman, in press) reveals the same pattern of results. Eight bets with varying stakes were used, with initial risk taking higher on low-stake bets. Across the eight bets the correlation between the mean of initial risk to a bet and mean risky shift elicited by that bet was .80. Recall that it was a comparable correlation obtained on choice-dilemmas items which inspired the group polarization hypothesis.

Various other risk measures have also been used. With an investment simulation task, Deets and Hoyt (1970) observed that average initial preferences were for securities toward the high-risk end of the continuum and that this tendency was strengthened by discussion. Runyan (1974) observed that subjects who advised another on risk taking initially elected more risk than did subjects who made binding decisions for another. This difference was slightly magnified by group interactions (p < .10). In a field experiment, McCauley, Stitt, Woods, and Lipton (1974) observed that groups were more cautious in race track betting than were individuals. The initial inclination of most bettors was also to prefer the favored horses and to avoid the long shots.

Finally, there are three small studies (Clement & Sullivan, 1970; Cohen & Ruis, 1974; Colligan & Giambra, Note 11) in which the investigators sought risky shift in classroom contexts in which students were asked, before and after discussion, to recommend one of several grading procedures (from quite risky—e.g., heavy weight on one final exam—to more conservative). In each of the three studies a shift to greater caution followed discussion, causing the investigators to question the reliability of risky shift or to label it an artifact of the laboratory. However, in two of these studies the investigators provided both initial means and shift means, and the findings are perfectly compatible with the group polarization hypothesis. For example, Clement and Sullivan observed that "the private choices before discussion were all from the conservative half of the eight choices presented (p. 244)." In the third study no pre-discussion data are presented.

All in all, it appears that recent demonstrations of risky and cautious shifts in risk behavior are quite compatible with the generalized group polarization hypothesis. In studies in which risky shift has not been observed, the results may be seen to confirm rather than refute the basic group locomotion phenomenon observed on choice-dilemmas items.

**Conclusion**

A review of recent literature on the effects of group discussion on various types of responses indicates that the evidence is generally consistent with the group polarization
hypothesis derived from research with choice-
dilemmas items. Group enhancement of ini-
tial mean tendencies is not always observed,
and when it does occur it is generally of
small absolute magnitude. A further difficulty
when trying to document the generality of
a phenomenon is the likelihood that negative
results have more often gone unreported
than have rejections of the null hypothesis
(Greenwald, 1975). But the trends are con-
sistent enough across a variety of task situa-
tions to conclude that some generality of
group polarization phenomenon has been es-
tablished. It appears that the conclusion of-
fered by Roseborough (1953) regarding the
state of knowledge on group problem solving
more than 20 years ago is equally appropriate
to the research literature on group polariza-
tion 20 years later:

We need not be further persuaded that group dis-
cussion processes have an effect on individual per-
formance even though there is a selective process
occurring in the reporting of studies. This proof has
only opened up new and troublesome problems con-
cerning the mechanisms by which this influence is
achieved and the conditions under which such an
empirical observation holds. (p. 279)

Theoretical Explanation of Group
Polarization

Having ascertained the empirical generality
of group polarization, we now consider possi-
bable explanations of it. What are the mecha-
nisms by which this group influence is
achieved and the precise conditions under
which the empirical observation holds? Our
goal is to understand the phenomenon in a
way that will account for the known condi-
tions under which group polarization occurs,
will predict the effects of other conditions not
yet studied, and will suggest when an en-
hancement of the initial mean tendency is not
to be expected. The initial mean tendency is
not causing the shift, it is only a crude index
to some actual group dynamics which do pro-
duce the effect. Thus, if we can identify the
general mechanism responsible for the en-
hancement of initial mean tendencies, perhaps
we can then use this understanding to better
predict the expected direction and extent of
group shift in a particular situation.

The findings we review in the remainder
of the paper are based mostly on choice-
dilemmas research. This operational consist-
ency has aided in the development of a well-
established set of empirical parameters. The
evidence reviewed earlier indicating that risky
shift is merely one subclass of a general
polarization phenomenon gives grounds for
also extending theory-testing research with
choice dilemmas to other situations in which
group polarization is observed. It is possible,
however, that the causal group dynamics will
vary somewhat across tasks. A good theory
will predict any situational differences.

The earlier reviews have already docu-
mented that group shifts do not appear to be
due, to any large extent, to the diffusion of
responsibility via the development of affective
bonds among group members, or to the
reduction of uncertainty resulting from mere
increased exposure to the stimulus materials,
or to a relationship between leadership ability
and propensity to take risks.1 We therefore
concentrate our analysis on three theories that
have recently received considerable attention.
Group decision rule explanations (e.g., major-
ity rule) predict shift by using statistical
schemes for combining individual preference
distributions into an expected group product.
According to one version of these statistical
explanations, no actual psychological changes
in preference need be postulated. Interper-
sonal comparison theories explain shift in
terms of social motivation. Subjects desire to
perceive and present themselves favorably, so
exposure to others' positions may stimulate
the subject to readjust his response in order
to maintain his image of social desirability.2

---

1 A potentially more viable notion than the gen-
eral leadership theory is that shift is due to dispro-
portionate influence by extreme group members. The
extremity of a subject's response to a specific item
may be related to the degree of his commitment,
awareness of arguments, and/or confidence. Avail-
able data do not strongly support this explanation
(see Davis, Kerr, Meek, & Rissman, in press; Myers
& Murdoch, 1972), although it is possible to incor-
porate extremity influence affects within the theories
discussed below.

2 Some have also called this value theory, a label
we avoid because of its ambiguity. All that the
interpersonal comparisons approach needs to presume
is that subjects enter into social interaction with
some shared initial suppositions as to what attitude
or course of action is desirable. It does not seem
necessary to debate whether, in a particular instance,
these suppositions also point to cultural values.
Informational influence explanations suggest that discussion generates arguments predominantly favoring the initially preferred alternative, including some persuasive arguments that the typical subject has not previously considered. Thus, a “message effect” evokes response change resulting from new cognitive learning.

Group Decision Rules

Several investigators have recently suggested that schemes for combining individual preferences might successfully predict the group product. Davis (1973) has enumerated several of these possible social decision schemes. The most popular has been the suggestion of majority rule (e.g., Burnstein, 1969; Cartwright, 1971; Lambert, 1969, 1971; Lamm, Trommsdorff, & Rost-Schaude, 1973). None of these authors contended that majority rule is the sole determinant of polarization, but they did convincingly point out the plausibility of this idea. A social decision scheme of majority rule predicts shift toward the dominant pole when the majority favors that direction and when there is skewness in the distribution of initial choices (which there will often be when the response mean departs from the midpoint of the scale). Perhaps when such a majority–minority constellation exists, deviant minority persons in the tail of a skewed distribution are moving toward agreement with the majority opinion, thereby creating a shift in the group mean. A statistical decision scheme such as this could potentially account for group polarization without recourse to postulating any real changes in individual preferences. (A decision scheme approach is not necessarily this indifferent to individual psychological change. Although one might consider shift a statistical artifact of a majority rule decision procedure, another view is to assume that majority influence produces genuine change in the minority.) A group decision scheme approach requires some initial diversity in choices and implies that the variance among choices will then generally be reduced after the discussion; and there is ample evidence that these conditions generally hold when shift is obtained (e.g., McCauley, 1972). Since the decision scheme approach is so intuitively compelling, it may be a disappointment that numerous other available data so sharply contradict it, at least as far as majority rule is concerned. We now enumerate in a roughly ascending order of significance those known facts which contradict the suggestion that group polarization is a statistical artifact of the application of social decision schemes. (Not all of these facts contradict a social influence interpretation of decision schemes.)

1. Group-induced shift appears to be internalized, not a temporary group product. The shift effect is about equally robust regardless of whether a group decision is required (Kogan & Wallach, 1967a; Lamm, 1967; Marquis, 1962; Wallach & Kogan, 1965; Wallach, Kogan, & Burt, 1968). Of course, when overt consensus is not required, an implicit consensus might nonetheless emerge. But it is furthermore true that nonparticipating observers of the group discussion also shift (e.g., Lamm, 1967) and that the group-induced changes in participant responses will persist for several weeks afterward (Johnston, 1968; Wallach, Kogan, & Bem, 1962—although this might be an after-effect of the consensus commitment).

2. When the opportunity for combining pretest decisions according to a decision rule is eliminated, shift still occurs. Shift effects have recently been obtained in numerous experiments using between-groups designs without pretest (e.g., Baron, Baron, & Roper, 1974; Davis et al., 1974; Gaskell, Thomas, & Farr, 1973; Myers & Bach, 1974; Myers, Bach, & Schreiber, 1974). Still, it might be contended that individual preferences are nonetheless expressed early in the discussion and then combined according to some implicit scheme. But when subjects discuss the substance of the items without an awareness of any response scale, shift still obtains (Myers, Bach, & Schreiber, 1974). This is also the case when subjects discuss a related response dimension (utilities of outcomes) and then respond on the probability scale they have not discussed (e.g., Burnstein, Miller, Vinokur, Katz, & Crowley, 1971). These are clear indications that genuine preference changes do occur, producing response change even in the absence of an opportunity for decision schemes to operate on the response dimension.
It is still possible, though, that a group might adjudicate their opinions (if not their responses) via some decision rule.

3. Skewness cannot account for group polarization. This is particularly relevant to the majority rule scheme, which depends on a skewed distribution of initial choices. On choice dilemmas, positively skewed distributions (i.e., with a risky majority) should produce risky shift, and negatively skewed distributions should yield a conservative shift. Several findings refute this prediction.

Vinokur (1969) reanalyzed raw data from earlier studies that had yielded risky shift and noted that negative skewness was equally as prevalent as positive skewness. Furthermore, both the positively and negatively skewed group distributions on the three most risky items produced equivalent shift.

Shifts in the group median, although slightly attenuated, are not significantly smaller than shifts in the group mean (Fraser, 1971; Myers & Aronson, 1972; Lamm et al., in press; Abend & Kogan, Note 12). This is contrary to any model that depends on skewness effects. For example, if the majority is really unaffected by the group discussion, then the median group member (who is one of the majority) should also be unaffected.

Group shift has also been shown to occur in dyads (although somewhat reduced), where obviously there can be no skewness in the initial responses (Baron, Baron, & Roper, 1974; Bateson, 1966; Lamm, Schauder, & Trommsdorff, 1971; Myers & Aronson, 1972; Swap & Miller, 1969). Most of the available group decision schemes would not predict polarization effects in dyads.

The majority rule model not only predicts a correlation between the initial mean of responses to an item and mean shift but, more specifically, also implies the same correlation at the level of specific group decisions (across groups within items). In other words, given a single item, the riskier a group's initial mean on that item, the more it should shift toward the dominant pole (because of increased likelihood that there will be a majority favoring that direction). Another group that on the same item initially tends in the opposite direction should then shift in that opposite direction. But in fact the correlation between initial mean and mean shift across groups within items is near zero (e.g., Teger & Pruitt, 1967). (Actually it is somewhat in the reverse direction, although ceiling and regression effects might contribute to this—a group initially near the extreme simply cannot move much further in that direction.) We return to this finding later.

Although group discussion generally results in decreased intragroup variance, as group decision schemes would predict, this may be a concomitant rather than an inherent feature of the shift phenomenon. By eliminating both pretesting and discussion of the response scale per se, Myers, Bach, and Schreiber (1974) demonstrated polarization without group convergence (as compared to variance in a control condition).

Finally, attempts to directly apply a majority rule model to discussion-produced shifts have produced only mixed results (e.g., Cvetkovich & Baumgardner, 1973; Moscovici & Zavalloni, 1969; Zaleska, 1976).

In summary, it is evident that while group decision models may be useful in other situations in which discussion is minimal or absent and the task is to reach agreement (e.g., Lambert, 1969), the models (or at least the majority rule model stressed in this analysis) are not a sufficient explanation of the group polarization findings we are seeking to explain. There are still a variety of other decision schemes that can be explored and with other specific tasks. But clearly, group-induced shift on choice dilemmas is something more than a statistical artifact.

We conclude with two general observations about prediction based on group decision models. First, the most stringent test of any model is not to predict the general trend of shift (since the general trend might occur for reasons irrelevant to the model), but rather to predict specific group outcomes. In specific group situations a given model might be quite an inaccurate predictor, but across groups the algebraic sum of its deviations may nonetheless be close to zero (because positive and negative errors cancel out, making the trend prediction accurate). In one study (Abend &
Kogan, Note 12) this in fact occurred. Their model was successful in predicting the general distribution of group products, but failed to pass the acid test of minimizing the sum of absolute deviations from actual choices. Thus, it is important that investigators who use this approach give their model the most rigorous possible test. (See, for example, Cartwright's, 1973, and Davis et al.'s, in press, analyses of percentage of "hits" in predicting specific group outcomes.)

Second, we suggest that even if statistical prediction by a model should prove to be quite accurate, this does not necessarily constitute a psychological explanation of the dynamics actually producing change in individual subjects. For example, it is conceivable (as Lamm, Trommsdorff, & Rost-Schaude, 1972, have pointed out) that a majority rule model might be successful for reasons implied by other available theories (e.g., Utterback, 1962, observed majorities to be more influential because in fact they possessed more cogent arguments). Put more positively, some social psychological explanations might be translatable into precise social decision scheme predictions (see Burns, 1967; Davis et al., in press). Thus, decision schemes need not be considered merely a statistical artifact approach to explaining group effects; they can, in fact, be a rigorous operational definition of certain psychological explanations.

Interpersonal Comparisons

A second class of explanations attributes group polarization to a type of normative social influence (connoting social-emotional processes and concern for favorable self-perception and self-presentation). More specifically, several versions of an interpersonal comparison approach suggest that mere exposure to the preferences of others is the necessary and sufficient condition for shift. Pruitt (1971a, 1971b) itemized the variations on this theme. In general, they suggest that a subject changes when he discovers that others share his inclinations more than he would have supposed, either because the group norm is discovered to be more in the preferred direction than previously imagined or because the subject is released to more strongly act out his preference after observing someone else who models it more extremely than himself. This theory, taken by itself, suggests that relevant new information which emerges during the discussion is of no consequence. Group polarization is a source effect, not a message effect.

A confusing array of reliable findings relevant to interpersonal comparison dynamics now exists. These diverse findings constitute a challenging intellectual puzzle to anyone who cares to search for a simple principle that will bring order to the mixture.

Differences between self, presumed other, and ideal scores. One well-known and widely substantiated assumption of the interpersonal comparisons approach is the observation from choice-dilemmas research that if, after responding, the subjects go back over the items and guess how their average peer would respond and then go back over the items a third time and indicate what response they would actually admire most, they tend to estimate the group norm as more neutral than their own initial response and their ideal as more extreme (see Myers, 1973, for a listing of more than 24 studies). Lamm et al. (1972) have also shown that not only do subjects indicate their ideal as more extreme than their actual response, but they also suspect that the same is true of their peers. The tendency of people to perceive themselves as more in what they consider to be the socially desirable direction than their average peer extends beyond the choice dilemmas (see Codol, Note 13). For example, most businessmen believe themselves to be more ethical than the average businessman (Baumhart, 1968), and there is evidence that people perceive their own views as less prejudiced than the norm of their community (Lenihan, Note 14).

On the choice-dilemmas task there are some puzzling order effects, however. The tendency to perceive others as "behind" oneself exists only when the self response is made prior to estimating the group norm (McCauley, Kogan, & Teger, 1971; Myers, 1974). Evidently it is after one has decided for himself that there is then a tendency to consider one's action as relatively admirable (by perceiving the average person as less admirable than oneself).
One version of interpersonal comparison theory suggests that in the group context, the subject typically discovers that the group norm is more supportive of his position than he had supposed (he is therefore outshining the other group members less than he had presumed). Consistent with this line of thinking, it has been shown that after discussion, subjects will indeed revise their estimates of the group norm more in line with reality, even when mention of initial responses has been suppressed (Clark et al., 1971; Ferguson & Vidmar, 1971; Myers, Bach, & Schreiber, 1974; Myers, Wong, & Murdoch, 1971).

Finally, it has been reliably demonstrated that subjects perceive other persons who have responded more extremely than themselves (in the direction of their ideal) as more socially desirable than persons who have not (Baron, Monson, & Baron, 1973; Jellison & Davis, 1973; Jellison & Riskind, 1970, 1971; Madaras & Bem, 1968). A parallel finding exists in the attitude literature (Eisinger & Mills, 1968): An extreme communicator on one's side of an issue tends to be perceived as more sincere and competent than a moderate.

While these findings are consistent with the interpersonal comparisons approach, their theoretical significance is somewhat ambiguous. For example, Burnstein, Vinokur, and Pichevin (1974) took an informational influence viewpoint and showed that people who adopt extreme choices are presumed to possess cogent arguments and are then presumably admired for their ability. They also demonstrated that subjects have much less confidence in others' choices than in their own, suggesting that the tendency to perceive others as more neutral than oneself simply reflects ignorance about others' choices. Given the ambiguous significance of these self–other–ideal difference scores, the important question becomes whether these difference scores predict response shifts, as interpersonal comparison theories would imply.

**Predicting shift with self, presumed other, and ideal difference scores.** This has been attempted in two different ways, correlationally and experimentally. First, we may inquire whether the perceived difference between oneself and others does in fact correlate with an individual's shift on a specific item. When a subject perceives himself as way ahead of his average peer, he should be more likely to suffer disconfirmation of his perceived relative position and so should be more stimulated to shift when informed of the actual group norm. But to the contrary, the self–other difference score on a particular item does not predict a subject's change following discussion of that item, even if his own initial response is partialed out (Lamm et al., 1971; Myers et al., 1971; Pruitt & Teger, 1967). Thus, a direct examination of interpersonal comparison theory at the level of particular decisions by individuals does not confirm the theory.

The experimental strategy is to compose groups on the basis of self versus perceived other difference or self versus ideal difference and observe whether this manipulation affects shift magnitude. Here the data are somewhat more comforting to interpersonal comparison theorists. Composing groups on the basis of self versus perceived other difference has yielded mixed results. Clark et al. (1971) observed increased risky shift when their subjects strongly underestimated peer risk acceptance, but Lamm et al. (1972) observed no effect of a similar manipulation. However, composing groups on the basis of the extent to which their ideal exceeds their actual response does affect shift in accord with interpersonal comparison theory (Lamm et al., 1971). The latter two studies by Lamm and his associates—combined with the observation that self–ideal difference scores are less affected by order of measurement than self versus perceived other differences (Myers, 1974)—suggest that the self–ideal discrepancy may be the more crucial element of a viable interpersonal comparisons approach. This conclusion is further reinforced by several independent observations of group-induced shift.

---

3 One other set of group composition experiments has grouped subjects high or low in social dependence (field dependence, need for approval, internal–external control, Machiavellianism). Subjects high in social dependence have not evidenced greater shift (Kogan & Wallach, 1967b; Lamm & Myers, in press; Lamm et al., 1971; Minton & Miller, 1970; Myers, Murdoch, & Smith, 1970; Wallach, Kogan, & Burt, 1967).
shift toward rather than away from the perceived norm (Baron, Roper, & Baron, 1974; Myers, Schreiber, & Viel, 1974; Vidmar, 1974). In two of the three studies it seems reasonable to presume that the subject's ideal may in fact have been in the same direction as the perceived norm but that an additional social constraint, surveillance by an adult who represented a conflicting norm, compromised the initial choices. The group interaction might then have liberated the subjects from the conflicting external norm.

Given these mixed results from attempts to predict shift with self-other-ideal difference scores, we now seek clarification of the role of interpersonal comparison dynamics in the final set of relevant studies. These attempts to provide the presumed necessary and sufficient conditions for shift—mere exposure to others' responses.

Choice shift following exposure to others' responses. One set of studies has manipulated the information about others' responses by providing fake norms. More than a dozen separate studies all show that subjects will move toward the manipulated norm (see Myers, 1973). This simply demonstrates that conformity effects can be demonstrated on choice-dilemmas items, as with numerous other measures. It is of interest, though, that the magnitude of the conformity effect is not generally increased much by making the fake norms consistent with the direction of one's ideals. Baron, Monson, and Baron (1973) recently demonstrated greater influence when the group deviated in the valued direction rather than the nonvalued direction. It is not certain whether this was due to the manipulation of confederate positions or to a concomitant increased potency of confederate information when arguing in the valued direction.

Other studies expose subjects to information about the actual initial choices of other subjects without any discussion or exchange of arguments. In general, the results from a series of such studies (see Myers, 1973, for bibliography) reveals either reduced but significant shift or no shift at all. Burnstein and Vinokur (1975) suggested that where an effect of mere exposure to others' choices was obtained, it may have been due to the stimulation this provided to think of arguments that others might have had for their choices. Consistent with this idea, they observed that exposure to others' choices produced shift only when subjects then wrote arguments on the item. If knowledge of others' choices was denied or if an opportunity to rethink the item was denied, no shift occurred.

On the other hand, it may be reasoned that in each of the studies producing minimal or nonexistent shift after exposure to others' attitudes, the subjects were first induced to bind themselves publicly to a pretest choice and then simply exposed to others' choices. It takes only a quick recall of some classic conformity studies (e.g., Asch, 1956) to realize that this was an excellent procedure for inhibiting response change. Perhaps discussion only serves the face-saving function of rationalizing response tendencies that are implicit after learning of others' preferences.

A recent experiment (Myers, Bach, & Schreiber, 1974) removed this problem by having the subjects respond to three risky choice dilemmas without a pretest commitment after merely being informed of the distribution of responses by 40 subjects in a control condition. While this provided a larger dose of interpersonal comparison than did the typical small group treatment, it maximized the possibility of a comparison effect uncontaminated by other variables. A surprisingly robust effect (increased risk taking) was evidenced when comparing these posttreatment responses with the control responses the subjects observed. This finding is counterintuitive in that it is the opposite of conformity—the subjects were fairly accurate in their guess of the average of the responses they had observed, and yet this exposure elicited a differentiation (polarization) from the observed norm.4

As we noted at the outset, the studies on interpersonal comparison dynamics do not lend themselves to a simple summary. However, enough evidence has been uncovered in

4 LeMaine (1974) has recently discussed the contribution to one's identity of differentiating oneself from others.
support of certain assumptions of the theory that we later attempt to integrate these assumptions into a theory of group-induced attitude change.

**Informational Influence**

The third major theoretical approach attributes the observed response changes to cognitive learning resulting from exposure to arguments during discussion. Discussion generates arguments predominantly favoring the initially preferred alternative, and for any given subject, some of these are likely to be persuasive new arguments. Thus, cognitive learning occurs mostly in the dominant direction and responses are modified accordingly. Each subject is processing and weighing information by a process akin to Anderson's (1971) information integration scheme, rather than comparing himself with the actual positions taken by others. The evidence relating to informational influence is largely of two types: experimental manipulations of the availability of arguments and content analyses of arguments generated by individuals and groups.

Several experiments have sought to stimulate argument exchange while eliminating information about others' choices. Normal shift effects result even when the subjects are prohibited from mentioning their initial choices (Clark et al., 1971; Myers et al., 1971; St. Jean, 1970) and even when the subjects exchange arguments without pretest and without awareness of the probability scale to which they will subsequently be exposed (Myers, Bach, & Schreiber, 1974). Similarly, discussing utilities of choice-dilemmas alternatives produces subsequent change in the probabilities chosen, even when probability choices are not directly discussed (and vice versa, discussing probabilities affects utility judgments—Burnstein et al., 1971; Vinokur, 1971a). While it is therefore clear that arguments have a persuasive impact, it is less clear whether this results from their teaching the subject something he did not know (informational influence) or from their telling the subject something about the general opinion positions of the persons who spoke them (interpersonal comparison). An ingenious experiment by Burnstein and Vinokur (1973), which provided arguments without clues as to the speakers' real attitudes (by keeping the subjects unaware of whether other group members were advocating their own initial choice or role playing support for a contrary one), indicates that informational influence is involved. In a recent attitude change study, Eagly (1974) has also experimentally demonstrated an effect of information comprehension. Studies of political attitudes in natural situations further evidence that increased information tends to polarize opinion (see Sears, 1969).

The second type of relevant evidence looks inside the "black box" of discussion to examine the actual distribution and nature of arguments expressed. It is now well documented that the direction of discussion arguments elicited by an item is an excellent predictor of the direction of shift generally obtained on that item (correlations in the area of .9 if a variety of items are used—e.g., Bishop & Myers, 1974; Ebbesen & Bowers, 1974; Morgan & Aram, 1975; Vinokur & Burnstein, 1974). These studies indicate that the initial mean of responses to an item predicts very well the trend of prediscussion and discussion arguments which in turn predict very well the mean shift. The predictive power of the initial response mean is thus captured almost entirely by the content of the subsequent discussion, suggesting that it is the nature of the expressed arguments which mediates the relationship between initial mean and mean shift.

Vinokur and Burnstein (1974, Note 15), Vinokur, Trope, and Burnstein (in press), and Bishop and Myers (1974) have formulated mathematical models of the presumed informational influence mechanisms. These models assume that the amount of group shift will be determined by three factors: the direction of each argument (which alternative it favors), the persuasiveness of each argument, and the originality of each argument (the extent to which it is not already known by the group members before discussion). In discussion, the potency of an argument will be zero if either the rated persuasiveness is zero (it is trivial or irrelevant) or if all group members considered the argument before discussion (in which case no cognitive learning will occur when the argument is heard). Since
the simple direction of arguments is such an excellent predictor of shift (without considering persuasiveness and originality), it is not easy to demonstrate the superiority of the models over a simple analysis of argument direction as undertaken by Ebbesen and Bowers (1974). Nonetheless, exhaustive analyses by Vinokur and Burnstein support their informational influence model by demonstrating that arguments do center on utilities of outcomes, that arguments consistent with the dominant tendency are rated as more persuasive than arguments opposing polarization, that the relevant arguments are indeed only partially shared prior to discussion, and that novel arguments have more impact than do unoriginal arguments. Furthermore, their model is even successful in predicting variation among group shifts within items. Recall that the prediscussion response means for particular groups do not account for variation among shifts within items. Earlier we noted that this latter finding was contrary to the majority rule model. It can, however, be accounted for by the informational influence assumption that there exists a pool of persuasive arguments for each item. A group that is fairly polarized on a particular item before discussion is presumably already in general possession of those arguments which polarize a group. A less extreme group has more to gain from the expression of partially shared persuasive arguments.

Two other sets of evidence are relevant to informational influence theory. First, experimentally manipulating the direction of the arguments to which subjects are exposed affects shift in accord with informational influence predictions (Baron, Dion, Baron, & Miller, 1971; Ebbesen & Bowers, 1974; Roberts & Castore, 1972; Silverthorne, 1971; Russo, Note 16). However, most of these studies were probably simultaneously manipulating information about others' positions, so they might also be interpreted as demonstrating interpersonal comparison or simple conformity effects.

While the above collection of findings constitutes compelling and consistent support for the causal role of cognitive learning in producing group polarization, the final set of findings suggests that cognitive learning is not a complete explanation by itself. Passive receipt of arguments outside an interactive discussion context generally produces reduced shift (e.g., Bishop & Myers, 1974; Burnstein & Vinokur, 1973; St. Jean, 1970; St. Jean & Percival, 1974). Likewise, listening to a group discussion generally elicits less shift than actual participation (e.g., Bell & Jamieson, 1970; Lamm, 1967; St. Jean, 1970). These findings—that active discussion generates more change than the passive receipt of information—probably should not have surprised us in light of the classic work by Lewin (1947) on the relative effects of discussion and lecture, educational research comparing the relative efficacy of information presentation versus participative discussion (McKeachie, 1968), and the observed ineffectiveness of information communications in field settings (Crawford, 1974). Studies by Moscovici, Doise, and Dulong (1972) and Moscovici and Lecuyer (1972) further indicate that instructions and seating patterns which interfere with the natural group interaction process reduce the shift effect.

A hint as to how the informational influence approach might be refined to accommodate these findings comes from theory and research on the role of cognitive learning and rehearsal in attitude change. McGuire (1972) pointed out that attention to and comprehension of arguments (cognitive learning) must be followed by conditions which also produce yielding in order for attitude change to be evidenced. Mere awareness of information is not a sufficient condition for attitude change. Consistent with McGuire's analysis, Greenwald (1968) observed that cognitive learning in a passive context was not sufficient to produce attitude change. Cognitive rehearsal of self-generated cognitive responses was also necessary for attitude change to occur. The subject must actively reformulate the information he has received in order for it to stimulate an internalization of attitude change.

As John Dewey (1900) once wrote, "A thought is not a thought unless it is one's own (p. 66)." It seems quite reasonable to presume that the social confrontation inherent in debate and discussion would motivate an active rehearsal process, even in those who are quietly contemplating their next remark.
Closely related to the rehearsal notion is the possibility of internalization resulting from the public verbal commitments that the subject makes during discussion. His verbal expressions are both one form of active rehearsal and a possible source of dissonance. Thus, attitude change may be motivated by a need for consistency with the verbal behavior elicited in discussion. Verbal commitment could produce the increased sense of involvement and certainty that Moscovici and Zavoloni (1969) believe to be inherent in group polarization. It should be noted that Ebbesen and Bowers (1974) did not observe that forced private generation of risky or cautious arguments affected shift. However, their procedure minimized both perceived choice and public exposure of the commitment, factors known to be important determinants of commitment effects (Kiesler, 1971). A more subtle procedure for manipulating the direction of rehearsed arguments did affect response shift (Burnstein & Vinokur, 1973).

In summary, the evidence for informational influence processes is compelling. However, it also appears that group polarization is not fully explained by a passive process of cognitive learning. More dynamic processes of cognitive rehearsal and verbal commitment are also likely contributors to the effects of talking in groups.

**Conceptual Integration**

We have seen that the isolation of important variables (i.e., arguments, interpersonal comparisons) demonstrates that each can have effects when other variables are held constant. But in reality the variables interact with and feed one another rather than functioning as separate, additive components of group influence. In life outside the laboratory, the expression of arguments and of one's position are often a part of the same communication sequence. Arguments convey implicit information about one's opinion. On the other hand, social motivations may influence the emission of arguments. Groups composed of subjects whose reported ideals exceed their actual responses are likely to select arguments which favor and reinforce their aspirations.

What we need is a conceptual integration which captures the interaction among these elements of social influence. Our proposal for a more holistic view of the group influence process (schematically represented in Figure 1) is inspired by the literature on social influence, which amply demonstrates that man is indeed both a social and a rational being, and by Kelman's (1974) analysis of the dynamics of attitude change. The scheme is simply an integrated summary of concepts established in past research on group shift. It is proposed as a net which can catch the diverse set of known facts about group shift and also place the group polarization concept within the scope of attitude theory—where it may profit from concepts developed in attitude research and in turn contribute to our understanding of how attitudes develop in their social context.

We assume that man's social motivation and rational faculties are dynamically interacting in social contexts. Figure 1 attempts to visualize this interaction and may be verbally summarized as follows. Social motivation produces a small direct impact by motivating the person to verbalize arguments that are socially desirable and in accord with his ideals. By offering arguments that tend toward the outer limits of his range of acceptability, the individual tests his ideals and also presents himself favorably to the group since, as we noted earlier, extremity in the direction of the ideal connotes knowledgeability and competence. (The expression of arguments is, of course, also determined by the individual's existing cognitions.) The expression of these arguments may have some direct impact on the relevant attitude, through a dissonance reduction or self-attribute process, and it also constitutes both a form of cognitive rehearsal for the speaker and of information to be received and responded to by the other group members. The resultant cognitive learning and rehearsal contributes significantly to subsequent attitude change. The strength of the evidence for informational influence effects is symbolized by the double arrows from the "cognitive foundation." Indeed, the evidence indicates that some polarization effect can result from a pooling of prediscussion arguments, apart from any added effects of social motivation.
To summarize, social motivation (a desire to perceive and present oneself favorably relative to others) may change an attitude through the interpersonal comparison process, but it also motivates the person to express socially desirable arguments. This verbal commitment may enhance the attitude, and it also serves a cognitive rehearsal function for the speaker and a cognitive learning function for the listeners. Thus, the effects of social motivation may be partly mediated by the learning and rehearsal that accompanies the hearing and speaking of discussion arguments.

There is one key assumption of this conceptual integration for which evidence has not yet been provided, namely, that discussion arguments are indeed influenced by the social context. Recent research provides such evidence. Arguments spoken in discussion more decisively favor the dominant alternative than do written arguments. This is indicated by evidence from Bishop & Myers [1974] and Ebbesen & Bowers [1974] and by a comparison of the data from Myers & Bishop [1971] and Silverthorne [1971] with the data from Burnstein & Vinokur [1973], Myers & Bach [1974], Stokes [1971], and Vinokur & Burstein [1974]. In writing an advisory brief or list of arguments, the subject is more impartial than when discussing the case with his peers. This might partly be due to a demand for balance in the instructions of the written brief conditions, but the finding is reliable enough to suggest the hypothesis that people are generally more polar in conversation than in writing. Perhaps this is because in conversation people are responding to other people, whereas when working alone the individual is responding only to the materials. A field demonstration observation of the effects of social context on communication content is Crawford’s (1974) observation of Catholic priests adapting their sermon arguments to fit the perceived attitudes of their parish. Laboratory research (e.g., Manis, Cornell, & Moore, 1974) also confirms that the average person will often ‘censor’ the information that he (or she) relays to others, so as to reduce the dissonance between the

---

5 The tendency for discussion arguments to be one-sided is probably not equal for all phases of a given discussion. Studies in speech-communications (see Fisher, 1974) suggest that one-sided discussion is especially likely after a choice direction has implicitly emerged and group members mutually reinforce their shared inclination.
The strength of the various vectors is expected to vary across situations. In more fact-oriented judgment tasks (group problem-solving tasks being the extreme case), the cognitive determinants will likely be paramount, although people will still be motivated to demonstrate their abilities. On matters of social preference, in which the social desirability of actions is more evident, the direct and indirect attitudinal effects of social motivation are likely to appear. The direct impact will occur in situations in which the individual has ideals that may be compromised by presumed norms but in which exposure to others' positions informs him that his ideals are shared more strongly or widely than he would have supposed. These situations—in which expressed ideals are a step ahead of prior responses—will also tend to elicit discussion content that is biased toward the ideals.

Informational influence mechanisms imply that shift will occur when discussion generates cognitive learning and rehearsal predominantly in one direction. Usually the mean of initial preferences is an index to the prevailing direction of cognitive learning, hence the group polarization phenomenon. But the mean initial preference is likely to be only a crude index of the predominant direction of cognitive change. For example, in the Myers and Bach (1974) study of separated pacifistic and militaristic groups, arguments generated by the two groups were similar, despite their initial differences in attitudes. In some other studies (e.g., some of the judgment and simulated jury experiments) the manipulation of the initial mean tendency may have been so unambiguous (e.g., by simple descriptive adjectives) as to preclude any persuasive new arguments relevant to the manipulation. It is also possible to conceive of situations in which a very familiar alternative, A, is initially favored over an unknown alternative, B, but shift occurs toward B because the potency of new information learned and rehearsed in support of B is greater than it is for A. Perhaps this is why Walker and Main (1973) observed bold new libertarian decisions in judicial groups even though individual judges did not make predominantly libertarian decisions when deciding alone.

**SIGNIFICANCE OF RESEARCH ON GROUP POLARIZATION**

A legitimate question raised by some who have assessed the risky-shift literature (Cartwright, 1973; Smith, 1972) is whether all the resources and time expended will in the end prove to have been worth it. With the addition of recent empirical and theoretical studies on group polarization, we again raise the question, What is being gained by the total research effort? The payoffs appear to be of two types: (a) knowledge directly gained about the effects of group discussion and (b) indirect heuristic payoffs in the form of new ideas about social influence and group process.

**Knowledge Gained About Group Effects**

The group polarization phenomenon is apparently a general outcome of group discussions of various sorts. The significance of this phenomenon may be evaluated in terms of the importance of small group communication for attitude change and in terms of possible pragmatic implications, as well as in terms of the actual experimental data summarized earlier.

What is the extent of small group influence on attitudes? McGuire (1969) noted, "It is clear that any impact that the mass media have on opinion is less than that produced by informal face-to-face communication of the person with his primary groups, his family, friends, co-workers, and neighbors (p. 231)." Possible reasons for the greater effect of face-to-face communication are summarized in Figure 1. In social interaction the target person is motivated to present himself favorably, and he is engaged in active cognitive rehearsal and verbal commitment. Thus, it is not surprising that in Western culture, group discussion seems increasingly integral to our social and organizational existence. Reliance on individual psychotherapy is being discarded in favor of mutual assistance in small groups (see Tyler, 1973), and group-centered organizational methods are displacing authoritarian management hierarchies (see Vroom, 1970, for a discussion of the impact of par-
The group polarization principle may help to explain what some observers feel are negative effects of group interaction (e.g., radical movements emerging from homogeneous subcultures and the difficulty of rehabilitation in closed environments such as penal and mental institutions), as well as effects that are often considered to be beneficial (e.g., in group therapy situations and in strengthened ethnic or religious identity subsequent to a grouping with similar others). Thus, we speculate that if the social and informational forces likely to be evoked by social interaction are in the desired direction, then discussion will likely be a useful strategy for producing more "effective" attitudes. If not, then the leader should either control the communication content so as to elicit desired arguments (as Lewin, 1947, did in some of his classic experiments on group discussion) or simply introduce the action as fait accompli (see Allport, 1954).

Some of the findings we have noted are experimental demonstrations of "groupthink" processes, which Janis (1972) has proposed to help explain decision fiascoes such as the failure to anticipate the Pearl Harbor attack, the invasions of North Korea and the Bay of Pigs, and the escalation of the Vietnam war. For example, his suggestion that group members "show interest in facts and opinions that support their initially preferred policy and take up time in their meetings to discuss them, but they tend to ignore facts and opinions that do not support their initially preferred policy" (p. 10) seems to be confirmed in the finding that discussion arguments are more in line with the dominant pole than are written arguments. Janis attributed this to a "suppression of deviant thoughts" resulting from "concurrence-seeking as a form of striving for mutual support" in order to "maintain self-esteem" (pp. 201-203).

This literature not only provides experimental confirmation for some of the groupthink dynamics proposed by Janis, it also suggests some additional ways to prevent them. The observations of comparison-induced polarization suggests that a group leader might profitably suppress mention of initial preferences while eliciting relevant arguments. The finding of greater polarity in discussion arguments than in written briefs suggests the usefulness of generating arguments by having participants individually write pro and con considerations for later sharing with the group. (The individual production of ideas has also been shown to generate a greater quantity of ideas than does group brainstorming—see the review by Lamm and Trommsdorff, 1973.)

One practical suggestion for doing this is the "nominal group" technique suggested by Van de Ven and Delbecq (1971):

Imagine a meeting room in which seven to ten individuals are sitting around a table in full view of each other. However, they are not speaking to each other. Instead, each individual is writing on a pad of paper in front of him. At the end of 10 to 20 minutes, a very structured sharing of ideas takes place. Each individual, in round-robin fashion, provides one idea from his private list which is written on a flip-chart by a recorder in full view of other members. There is still no discussion, only the recording of privately generated ideas. This round-robin listing continues until each member indicates that he has no further ideas to share. . . . Generally, a spontaneous discussion then follows for a period (in the same fashion as an interactive group meeting) before nominal voting. Nominal voting simply means that the selection of priorities . . . is done by each individual privately, and the group decision is the pooled outcome of the individual votes. (p. 204)

Heuristic Payoffs

Pruitt (Note 18) regarded research on group shift primarily "as a heuristic procedure for generating a new theoretical paradigm in the general area of group locomotion (p. 3)." He suggested that the elucidation of a convincing explanation of the shift effect may enliven the field of group dynamics "as people begin to develop and test the implications of the theoretical assumptions underlying this explanation (p. 3)."

Very recently investigators in this field have begun to cease thinking of the group change effect as risky shift and are instead using the paradigm to develop basic knowledge about group process and about the dynamics of attitude change in social situations.
(as contrasted with research on passive recipients receiving programmed communications). Hopefully, theoretical ideas developed to explain group polarization will be useful beyond this paradigm.

For example, interpersonal comparison theory stimulates questions about the attitudinal effect of exposure to others' attitudes. Might there be conditions under which social contagion effects result from feedback concerning others' attitudes? Brown (1974) stated the generalized hypothesis clearly:

The suggestion is that just as there are many situations in which individuals will be motivated to conform, to huddle up close to the central tendency (as in perceptual judgments), there is another class of situations in which people will be motivated to fall on one or the other side of the central tendency because they seek not to be average but rather better than average, or virtuous. To be virtuous, in any of an indefinite number of dimensions, is to be different from the mean—in the right direction and to the right degree. (p. 469)

For example, if an opinion poll reveals that others support your inclinations to an extent that surprises you, does this tend to reinforce your opinion on the matter? There has been a substantial amount of research on the effects of misinforming a subject concerning others' attitudes and behavior, but almost no research in social psychology of public opinion using a paradigm that more closely approximates social reality by giving accurate information concerning others' attitudes. There is known to be no general "bandwagon effect" resulting from the publication of political polls. But interpersonal comparison theory provides a more precise prediction: Attitudes will change if and when there is a discrepancy between one's social perception of the opinions of his reference group and his revised perception after exposure to the actual distribution of opinions.

Of course, we are observing others' positions in many ways other than opinion polls. Alker (Note 19) suggested that recent increases in political attitude polarization noted in national polls "may have been facilitated by repeated exposure in the media of persons holding extreme views." This suggestion is compatible with Pruitt's (1971a, 1971b) "release theory" of interpersonal comparison effects. The media present models who embody their viewpoint with great clarity, consistency, and courage. We have noted earlier the admiration which is accorded extreme representatives of one's viewpoint. This may therefore release a sympathizer to then align himself with the more extreme forms of his own opinion.

In summary, research on group polarization has provided new knowledge about the outcomes of group discussion, and the need to explain the phenomenon has stimulated new concepts which may enrich our understanding of attitude change in social situations. If further research on these concepts should prove to be fruitful, then we may console ourselves that the effort represented by more than 300 manuscripts since Stoner's (1961) discovery of the risky shift has indeed been of value.

REFERENCE NOTES
presented at the meeting of the Southeastern Psychological Association, April 1972.


REFERENCES


Asch, S. E. Studies of independence and conformity: A minority of one against a unanimous majority. Psychological Monographs, 1956, 7(9, Whole No. 436).


Baron, R. S., Roper, G., & Baron, P. H. Group discussion and the stingy shift. Journal of Personality and Social Psychology, 1974, 30, 538-545.


Bell, P. R., & Jamieson, B. D. Publicity of initial decisions and the risky shift phenomenon. Journal of Experimental Social Psychology, 1970, 6, 329-345.

Bishop, G. D., & Myers, D. G. Informational influence in group discussion. Organizational Behavior and Human Performance, 1974, 12, 92-104.


Burnstein, E., & Vinokur, A. What a person thinks upon learning he has chosen differently from others: Nice evidence for the persuasive-arguments explanation of choice shifts. Journal of Experimental Social Psychology, 1975, 11, 412-426.


Doise, W. Intergroup relations and polarization of individual and collective judgments. *Journal of Personality and Social Psychology*, 1969, 12, 136–143. (b)


Kelman, H. C. Attitudes are alive and well and gainfully employed in the sphere of action. *American Psychologist,* 1974, 29, 310-324.


McGuire, W. J. The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.),


Vinokur, A. A review and theoretical analysis of the effects of group processes upon individual and group decisions involving risk. *Psychological Bulletin, 1971, 78*, 231–250. (b)


(Received February 28, 1975)