UC Santa Barbara
UC Santa Barbara Previously Published Works

Title
Choice Shift and Group Polarization

Permalink

Journal
American Sociological Review, 64(6)

ISSN

Author
Friedkin, Noah E

Publication Date
1999-12-01

DOI

Peer reviewed

CHOICE SHIFT AND GROUP POLARIZATION[*]

Noah E. Friedkin
University of California at Santa Barbara

I extend the theoretical domain of sociology into an area of social psychology that heretofore has been the exclusive domain of psychologists. Specifically, I develop a social structural perspective on the choice shifts that individuals make within groups. During interpersonal discussions of issues, choice shifts occur when there is a difference between group members' mean final opinion and their mean initial opinion. Explanations of choice shifts have emphasized group-level conditions (e.g., a norm, a decision rule, a pool of persuasive arguments, a distribution of initial opinions). I argue that choice shifts are a ubiquitous product of the inequalities of interpersonal influence that emerge during discussions of issues. Hence, I bring choice shifts squarely into the domain of a structural social psychology that attends to the composition of networks of interpersonal influence and into broader sociological perspectives concerned with the formation of status structures.

The etiology of networks of interpersonal interaction and the effects of these networks on actors' agreements and decisions have been longstanding concerns of sociologists. There was a time when psychologists, following the lead of Kurt Lewin, also were active in studying the origins and effects of interpersonal networks (Cartwright and Zander 1968; Festinger, Schachter, and Back 1950; Newcomb 1961). As an area of study within psychology, however, the field of group dynamics has declined, and although the concepts of group membership and group effects have been retained, psychologists now rarely grapple with the network structure of groups, the social processes that unfold in these networks, and the contributions of these network structures and processes to individual and collective outcomes. It is sociologists—concerned with the origins of influence networks and status structures, the effects of social exchange networks on actors' bargaining behavior, and the effects of social networks on information flows, consensus formation, and collective action—who continue to grapple with these matters.

With few exceptions, however, sociologists have restricted their studies to questions that have not been pursued by psychologists (Cook, Fine, and House 1995; Lawler, Ridgeway, and Markovsky 1993). This "boundary" is strange and artificial because some questions that have been addressed by psychologists touch on themes within the sociological imagination. Does a group discussion of an issue foster extreme or risky decisions (group polarization)? What conditions in a group lead to a "loss" of self (deindividuation) and a high susceptibility to interpersonal influence (pressures toward uniformity)? Under what conditions do minority opinions in a group influence the opinions and decisions of group members who are in the majority? There is every reason why sociologists should take up such questions and bring our structural perspectives to bear.

In this article, I develop a social structural perspective on one of these questions—the explanation of choice shift and group polarization in small groups. The study of choice shift and group polarization is a prominent

line of work in social psychology that has been developed mainly by psychologists. The research is not well-known among sociologists, although it deals with a central theoretical interest of sociologists: the effect of group membership on individual attitudes. A choice shift is said to occur when, after a group's interaction on an issue, the mean final opinion of group members differs from the members' mean initial opinion. Group polarization is said to occur when the choice shift is in the same direction as the mean initial opinion (e.g., if on some issue, the initial attitude of the average member is positive [negative], then the subsequent attitude of the average member after group discussion will be more positive [negative]). An explanation of choice shift is fundamental because it would also explain group polarization. Group polarization always involves a choice shift, but a choice shift can occur that does not entail group polarization (i.e., a choice shift that is in the opposite direction of the initial inclination of the group).[1]

Research on choice shifts originated with a finding about choice dilemmas (issues in which a level of acceptable risk on a course of action is being debated) that indicated that the decisions of groups involve higher levels of risk-taking than do the decisions of individuals (Stoner 1961). This finding, known as the "risky shift," stimulated a huge number of studies. Recently, research on choice shifts has developed as a counterpoint to classical models of influence, which emphasized mechanisms of accommodation and conformity (Allport 1924; Asch 1951; Festinger et al. 1950; Sherif 1936). The assumption motivating this recent work is that the ubiquitous occurrence of choice shifts indicates a serious limitation of the classical models of social influence and demands the development of new theories based on different assumptions (McGarty et al. 1992; Turner and Oakes 1989).

I develop an account of choice shifts that was suggested by Cartwright (1971) in his analysis of the "risky shift" literature. Cartwright pressed for an analysis of the individual-level process that produces a choice

shift. He suggested that a choice shift could be produced by the process of interpersonal accommodation itself (i.e., that no new process or separate group effect produces these shifts), and he pointed to the formal theory of social power developed by French (1956) and Harary (1959) as a possible starting point for such an explanation. However, Cartwright stopped short of demonstrating that an "Ockham's razor" principle might be applied. Much of the subsequent work on choice shifts has continued to argue that a mechanism of social influence only predicts a convergence of opinion to the mean of initial opinions and, therefore, cannot account for the choice shift phenomenon.

I revisit Cartwright's (1971) arguments, applying a network theory of social influence I developed with Johnsen (Friedkin 1991, 1998; Friedkin and Johnsen 1990, 1997), which stems from the early work of French (1956) and Harary (1959). This theory potentially has a broad range of applications to problems in collective behavior, organizational sociology, political sociology, and small group sociology (Marsden and Friedkin 1994). I apply the theory to choice shifts and polarization in small groups and show how it can inform and integrate previous work on these phenomena. With this theory, I demonstrate (a) that choice shifts can arise simply from inequalities in the relative influence of persons in a group interaction, (b) that such inequalities are part of the process of interpersonal influence, and (c) that group polarization is not a ubiquitous, independent feature of the opinion formation process in group discussions of issues. In effect, I argue that a choice shift is the product of the group's social structure in which certain members have more influence than others during the opinion formation process.[2] This theoretical approach encompasses, as a special case, the classical work on pressures toward uniformity which show that such

---

[1] Choice shift and group polarization have been defined and measured in different ways. I review these alternative approaches later in this article.

[2] Thus, sociological work on status structures in small groups can have an important bearing on choice shifts (Berger, Fisek, et al. 1977; Berger, Ridgeway, et al. 1998; Ridgeway et al. 1998). Status differentials foster unequal interpersonal influences that are a precondition of choice shifts. I am not concerned with this mechanism here, although it suggests a theoretical link between work on status-organizing processes and choice shifts.

group pressures may lead to a convergence of opinion toward the mean of the initial opinions of the members of a group.

I also open a pathway for sociological exploration of a domain—small group dynamics—that traditionally has been a province of psychology. The vehicle for this exploration, a network theory of social influence, does not overturn extant work on group dynamics but instead envelops and enhances this extant work by attending to the *structure* in which interpersonal processes unfold in a group. The successful integration of network theory into the study of group dynamics should provide a better theoretical understanding of those groups in which more or less elaborate social structures affect individual and group outcomes, and should yield more precise predictions of these outcomes.

THEORETICAL BACKGROUND

The many studies of choice shifts have discredited the idea that group discussion produces greater acceptance of risk. The more general phenomenon of choice shifts has been thoroughly documented, however, and continues to motivate research (Cartwright 1971, 1973; Clark 1971; Dion, Baron, and Miller 1970; Isenberg 1986; Lamm and Myers 1978; Pruitt 1971; Vinokur 1971; also see the more recent work cited below). Currently, there are four main approaches to the study of choice shifts—*social comparison theory, persuasive arguments theory, self-categorization theory*, and *social decision scheme theory*. These approaches do not describe how the influence network of groups affects choice shifts, although they are consistent with the occurrence of this mechanism. I suggest that social influence network theory not only integrates important features of these approaches, but also provides a more concrete model of the social process by which choice shifts are produced via interpersonal influences.

Social Comparison Theory

According to social comparison theory (Baron et al. 1996; Goethals and Zanna 1979; Sanders and Baron 1977), actors initially espouse opinions that are less extreme than their true opinions because of a fear of being labeled deviant and that (after group discussion reveals that others espouse similar but more extreme opinions) they shift their opinions toward their true values and may even leapfrog over the more extreme opinions of others:

> This realization either "releases" the moderate members from their fear of appearing extreme, or motivates moderate members to "compete" with the extreme members to see who can come closest to espousing the most admired position. In either case, the moderates are motivated to adopt more extreme positions, while there is no corresponding pressure on extreme members to moderate their opinions (although, of course, simple conformity pressure may lead to some small amount of moderation by extreme members). The net result is an overall polarization of opinions, that is, a choice shift. (Sanders and Baron 1977:304)

Group discussion, by exposing group members' opinions, erodes moderate opinions and reinforces the normative attraction of adopting extreme positions. This theory does not address any of Festinger's (1954) hypotheses about the effect of an opinion discrepancy on interpersonal influence, nor does it draw on the idea of a structure of interpersonal influences. Instead, the theory proposes a process of de-inhibition followed by status competition, which contributes to the development of opinions in a group—a process that is *separate* from the action of interpersonal influences on the reduction of opinion differences.

Persuasive Arguments Theory

Persuasive arguments theory (Burnstein 1982; Burnstein and Vinokur 1973, 1977) explains group polarization on the basis of the *content* of the arguments that arise during a discussion. The hypothesis is that there is a pool of arguments that could be applied to any issue, and that discussants draw those arguments from the pool which support their initial opinions. Hence, if the distribution of initial opinions is biased in a particular direction, the sample of arguments expressed will reflect this bias and thus will influence the direction of a choice shift:

> When the preponderance of arguments in the pool favors a particular alternative, the average prior attitude reflects the direction and magnitude of this preponderance. Further thought or

discussion leads to polarization toward the alternative that initially elicits more and/or better arguments. (Burnstein and Vinokur 1977:316)

This argument has been refined in various ways (Lamm and Myers 1978), but the key idea is that choice shift and group polarization depend on the actual or implicit *arguments* for positions that result from group discussions or a rethinking of the issue. This theory does not address the structure of interpersonal influences that is based on the particular arguments and responses of group members to these arguments. Instead, it emphasizes the effects of the *set* of arguments that emerge on an issue. Nevertheless, the theory is not inconsistent with an analysis of individual differences in the pattern of persuasive argumentation in a group.

Self-Categorization Theory

Self-categorization theory (Mackie 1986; McGarty et al. 1992; Turner 1985; Turner and Oakes 1989) explains group polarization on the basis of the actors' conformity to an extreme norm or prototypical position of the group. Polarization may occur after group members have defined the social identity of the group and identified with the group. Given these identifications, group members modify their opinions in the direction of the normative opinion:

> The basic argument of [self-categorization] theory relevant to polarization is that people are conforming to a shared in-group norm, but that the norm is not the pre-test average but rather the position most *prototypical* of the group. The prototypicality of in-group members is defined by means of the *meta-contrast principle* . . . : the less a person differs from in-group members and the more he or she differs from out-group members, the more representative is he or she of the in-group. Thus the prototype is the position which best defines what the group has in common *in contrast to other relevant out-groups*. . . . The most consensual, normative position is understood both as a defining categorical property of the group prior to interaction and as the position on which members converge through social interaction. Both convergence on the mean and polarization are explained as the result of members' moving towards the consensual position of their group. (McGarty et al. 1992:3, emphasis in original)

The out-group positions that enter into the definition of the prototypical in-group position are all those discrete positions that are *not* endorsed by at least one group member. Hence, self-categorization theory permits the prototypical or normative opinion to be *any* position in the range of group members' initial opinions (e.g., a prototypical position could be the group's mean initial opinion or it could be one of the group's most extreme opinions depending on the out-group frame of reference).

Proponents of self-categorization theory argue that group members modify their opinions to reduce the discrepancy between their initial opinions and a group norm because they identify with the group norm and because they find members' opinions more or less persuasive depending on the extent to which they represent the group norm. Thus, Turner (1987) states, "the informational value of a response or its 'persuasiveness' is *exactly equivalent* to the degree that it is perceived to be exemplary of some in-group norm or consensus" (p. 154). The influence network of a group is specified completely by the distances between members' positions and the prototypical position.

Social Decision Schemes

Social decision scheme theory postulates that group decisions can be understood in terms of the initial distribution of opinions in the group and a decision scheme, or decision rule, that members use to obtain a decision (Davis 1973; Kerr 1981; Stasser, Kerr, and Davis 1989). A decision scheme may be based on majority rule, the arithmetic mean of initial opinions, the median of initial opinions, the most extreme initial position, and so on (Laughlin 1980; Zuber, Crott, and Werner 1992).

Social decision scheme theory relies on the distribution of initial opinions to specify the relative influence of the alternative initial positions of group members on an issue. Although the alternative initial positions on an issue can differ in their influence (depending, for example, on the number of persons who hold a particular position), individual differences in interpersonal influence are not taken into account:

Clearly, in most groups there are some members who are more persuasive, knowledgeable, intimidating, indomitable, or able than others; in fact, some of the SDS [social decision scheme] research . . . has empirically made this point. SDS and STS [social transition scheme] models can be modified to incorporate individual differences, but this involves considerable complexity. (Kerr 1992:70)

Social influence network theory (below), which explicitly incorporates such individual differences into the process of opinion change, is entirely consistent with social decision scheme theory. In fact, it could be described simply as a social combinatorial model of weighted averaging. In social influence network theory, the relative weights of particular opinions emerge from the process of interpersonal influence, and when a consensus is reached, one can derive the main outcome of social decision scheme theory—the relative contribution of particular positions on an issue to a group decision.

SOCIAL INFLUENCE NETWORK THEORY

Social influence network theory has been under development by social psychologists and mathematicians since the 1950s (DeGroot 1974; French 1956; Friedkin and Johnsen 1990; Harary 1959). This theory originated with French (1956), who formed a simple model of how a network of interpersonal influence affects opinion formation. French's model was given a more general form by Harary (1959) and DeGroot (1974). While these initial formulations described the origins of group consensus, they did not explain influence processes that reduced opinion differences short of a consensus. Friedkin and Johnsen's (1990) work addressed this limitation (i.e., accounted for the emergence of settled patterns of disagreement in a group) and formally integrated this line of theory with extant sociological approaches to network effects (Burt 1982; Doreian 1981; Erbring and Young 1979; Friedkin 1990; Marsden and Friedkin 1994).

The distinguishing characteristic of this theoretical approach is its attempt to model the flows of interpersonal influence that affect actors' opinions; compare the work on social decision schemes (Stasser et al. 1989)

and social impact theory (Latane 1981). These flows of influence arise from actors' repeated efforts to reduce differences of opinion in the context of an influence structure in which interpersonal influences of group members can vary in strength. The process of opinion formation can rarely be reduced to accepting or rejecting the consensus of others; typically, individuals form their opinions in a complex interpersonal environment in which influential opinions are in disagreement and liable to change. How opinions and consensus form in such complex circumstances is the subject of the theory.

Social influence network theory postulates a simple recursive definition for the influence process in a group of N actors:

$$\mathbf{y}^{(t)} = \mathbf{AWy}^{(t-1)} + (\mathbf{I} - \mathbf{A})\mathbf{y}^{(1)}, \qquad (1)$$

for $t = 2, 3, \ldots$, where $\mathbf{y}^{(t)}$ is an $N \times 1$ vector of actors' opinions on an issue at time t, $\mathbf{W} = [w_{ij}]$ is an $N \times N$ matrix of interpersonal influences, $\left(0 \le w_{ij} \le 1, \sum_j^N w_{ij} = 1\right)$, and $\mathbf{A} = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{NN})$ is an $N \times N$ diagonal matrix of actors' susceptibilities to interpersonal influence on the issue $(0 \le a_{ii} \le 1)$. This model stipulates that actors modify their opinions by forming a weighted average of influential opinions (Anderson 1996; Anderson and Graesser 1976; Graesser 1991). The group's influence network, $\mathbf{W}$, describes the pattern and magnitude of these direct endogenous interpersonal responses. Actors are influenced not only endogenously by the opinions of other actors, however, but are also influenced exogenously, at each point in the process, by the conditions that formed their *initial* opinions. The balance-of-forces (relative weight) of the endogenous and exogenous influences for each actor is described by $\mathbf{A} = [a_{ij}]$, the coefficients of social influence.

I apply a version of this general formulation (Friedkin 1998), which stipulates that

$$\mathbf{W} = \mathbf{AC} + \mathbf{I} - \mathbf{A}, \qquad (2)$$

where $\mathbf{C} = [c_{ij}]$ is an $N \times N$ matrix of relative interpersonal influences $\left(0 \le c_{ij} \le 1, c_{ii} = 0, \sum_j^N c_{ij} = 1\right)$.[3] This formulation equates ac-

---

[3] An even distribution of interpersonal influ-

tors' lack of susceptibility to interpersonal influence with the weight that they place on their own initial opinions,

$$w_{ii} = 1 - a_{ii}, \qquad (3)$$

and distributes the cumulative weight of others $\left(a_{ii} = \sum_j w_{ij}, j \neq i\right)$ according to the relative strength of the interpersonal influences, that is, for $i \neq j$,

$$w_{ij} = a_{ii}c_{ij}. \qquad (4)$$

Assuming the process reaches an equilibrium, such that the process is no longer inducing opinions changes, equation 1 is

$$\mathbf{y}^{(\infty)} = \mathbf{AW}\mathbf{y}^{(\infty)} + (\mathbf{I} - \mathbf{A})\mathbf{y}^{(1)}, \qquad (5)$$

and actors' settled opinions are

$$\mathbf{y}^{(\infty)} = \mathbf{V}\mathbf{y}^{(1)}, \qquad (6)$$

where

$$\mathbf{V} = \lim_{t \to \infty} \left[(\mathbf{AW})^{t-1} + \left\{ (\mathbf{AW})^{t-2} + (\mathbf{AW})^{t-3} \right. \right.$$
$$\left. \left. + \cdots + (\mathbf{AW}) + \mathbf{I} \right\}(\mathbf{I} - \mathbf{A}) \right] \qquad (7)$$

is a matrix of reduced-form coefficients describing the total interpersonal effects that transform initial opinions into final opinions. The coefficients in $\mathbf{V} = [v_{ij}]$ are nonnegative $(0 \le v_{ij} \le 1)$, and each row of $\mathbf{V}$ sums to unity $\left(\sum_j v_{ij} = 1 \right)$. Hence, v_{ij} gives the *relative weight* of the initial opinion of actor j in determining the final opinion of actor i for all i and j. If $(\mathbf{I} - \mathbf{AW})$ is nonsingular, then $\mathbf{V}$ can be derived directly from

$$\mathbf{V} = (\mathbf{I} - \mathbf{AW})^{-1}(\mathbf{I} - \mathbf{A}). \qquad (8)$$

Otherwise, $\mathbf{V}$ can be obtained numerically from equation 7.

When a consensus is formed in a group, $\mathbf{V}$ will have the following form,

$$\begin{bmatrix} v_{11} & v_{22} & \cdots & v_{NN} \\ v_{11} & v_{22} & \cdots & v_{NN} \\ \vdots & \vdots & \vdots & \vdots \\ v_{11} & v_{22} & \cdots & v_{NN} \end{bmatrix},$$

in which each actor's initial opinion makes a particular relative contribution to the emer-

gent consensus. Hence, the total effect of each alternative initial position can be aggregated (i.e., the total effects of the persons who hold a particular position can be summed) to obtain the relative weights of the various issue positions in determining the group outcome. For example, if actors 1, 3, and 5 held the same initial opinion on an issue, then the total weight of that issue position in determining the group consensus is $v_{11} + v_{33} + v_{55}$. The derivation of these weights shows that this approach achieves the same end as social decision scheme theory while taking into account individual differences in interpersonal influence.

Social influence network theory also brings together certain psychologists' views of interpersonal influence as a strain toward the mean of actors' initial opinions and sociologists' views of such influence as a source of inequality and domination. These outcomes—convergence on the mean of initial opinions and convergence on the initial opinion of a particular actor or subgroup—are special cases of the theory. Because the social structure of groups varies, the group outcomes appear in a variety of forms: Opinions may settle on the mean of group members' initial opinions; they may settle on a compromise opinion that differs from the mean of initial opinions; they may settle on an initial opinion of a group member; or they may settle on altered opinions that do not form a consensus. Because these are common outcomes, a general model of social influence must encompass them all, and the applicability of the present theory to these different situations is part of its appeal.

Social influence network theory integrates the persuasive arguments, social comparison, and self-categorization approaches to the extent that the effects of an argument, social comparison, and prototypical position can be represented directly as interpersonal influences, or as mediated by such influences. Group norms or other group-level conditions can affect opinions via an effect on persons' initial opinions, their susceptibilities to interpersonal influence, or their own interpersonal influence on others. The interpersonal influences described by the influence network, $\mathbf{W}$, can be based on the visibility (mere awareness) of an opinion difference, persuasive argumentation, or group members' re-

---

ences for actor i (i.e., $c_{ij} = 1/(N-1)$ for all j) is assigned in the special case when actor i is not influenced by any other actor.

sponses to status characteristics and power bases (e.g., expertise, reward power, authority, etc.).

An example from one of the experiments to be reported illustrates how this theory bears on choice shifts. This experimental group dealt with a monetary reward issue on which opinions (expressed in thousands of dollars) were formed. The initial opinions and final consensus of the group were, respectively,

$$\mathbf{y}^{(1)} = \begin{bmatrix} 1000 \\ 500 \\ 1000 \\ 1000 \end{bmatrix} \quad \mathbf{y}^{(\infty)} = \begin{bmatrix} 500 \\ 500 \\ 500 \\ 500 \end{bmatrix}.$$

Clearly, the choice shift in this case was substantial: The mean of the initial opinions is $\bar{y}^{(1)} = 875$, the mean of the final opinions is $\bar{y}^{(\infty)} = 500$, and the choice shift $\left(\bar{y}^{(\infty)} - \bar{y}^{(1)} = -375 \right)$ indicates that the group discussion lowered the average reward. The influence network of the group,

$$\mathbf{W} = \begin{bmatrix} 0 & .78 & .11 & .11 \\ 0 & 1 & 0 & 0 \\ .14 & .53 & 0 & .33 \\ 0 & .50 & .50 & 0 \end{bmatrix},$$

describes the distribution of relative interpersonal influences (the derivation of this network is described later). The main diagonal of $\mathbf{W}$ shows the actors' self-weights ($w_{ii} = 1 - a_{ii}$), and therefore $\mathbf{A} = \mathrm{diag}(1, 0, 1, 1)$ describes the actors' susceptibilities to interpersonal influence. The off-diagonal entries of $\mathbf{W}$ are the interpersonal influences. For example, $w_{12} = .78$ indicates that the direct (unmediated) relative influence of actor 2 on actor 1 is .78. The total effects matrix,

$$\mathbf{V} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

indicates that actor 2 dominated the group outcome, even though this actor held an initial opinion that deviated from an initial agreement of the other three members. In this case, the predicted final opinions,

$$\hat{\mathbf{y}}^{(\infty)} = \mathbf{V}\mathbf{y}^{(1)} = \begin{bmatrix} 500 \\ 500 \\ 500 \\ 500 \end{bmatrix},$$

correspond exactly to the observed final opinions. This example shows that the theory is consistent with minority influence. Other outcomes (disagreement, final consensus on the mean of initial opinions, group polarization, conformity of a deviant opinion to an initial agreement among the other members of a group) also emerge as special cases of a group's distribution of initial opinions and influence network.

Social influence network theory does not support a reductionist argument that collective outcomes (such as consensus or disagreement) can be understood simply on the basis of individuals' independent responses to their situations. Although the theory rests on a model of how individuals cognitively integrate conflicting opinions, the outcome of this process depends on (and cannot be understood apart from) the social structure in which the process occurs. This social structure consists of the particular configuration of members' attributes (initial preferences, susceptibilities to influence) and interpersonal relations. A change in the configuration of these group attributes and relations may produce a substantial change in individual and collective outcomes. In this way, groups can be said to have effects on their members.

METHODS

Three Experiments

I draw upon three experiments involving groups of different sizes: dyads, triads, and tetrads. The form of these experiments was roughly the same: (1) Each member of a group privately recorded their initial opinion on an issue; (2) the members of the group then discussed the issue over a telephone network in which certain pairs of members were allowed to communicate but in which all members were aware of the overall pattern of communication channels; (3) after some specified time, or upon reaching group consensus or a deadlock, group members privately recorded their final opinions on the is-
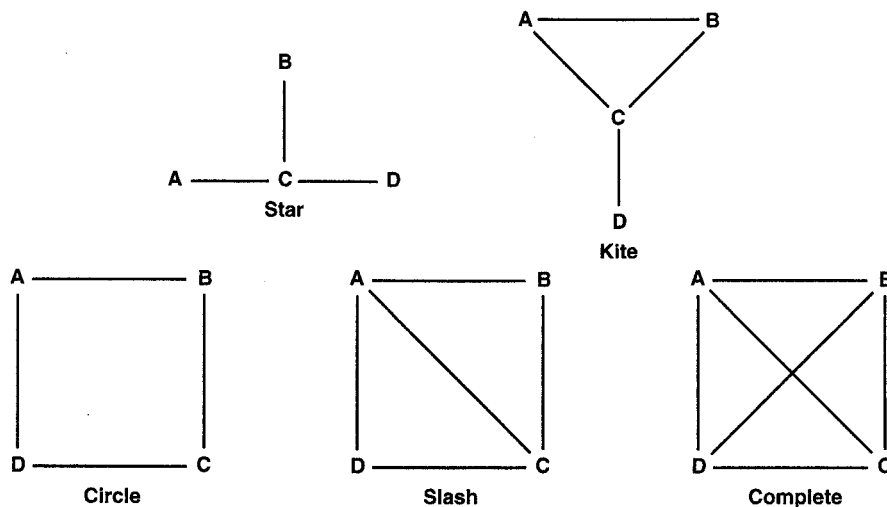
Figure 1. Network Structures Used in the Study

Note: Illustrations shown are for four-person groups.

sue and provided estimates of the relative interpersonal influences of the other group members on their final opinions. These experiments provided data on three of the four theoretical constructs involved in the influence model described by equation 5: actors' initial opinions, $y^{(1)}$; actors' final opinions, $y^{(\infty)}$; and actors' relative interpersonal influences, C, on an issue. Given these data, the fourth construct, actors' susceptibilities to interpersonal influence, A, were estimated using the method described in Appendix A.

These experiments were developed to test a larger set of theoretical questions than I address here and are part of a larger project on the influence structures of small groups. Hence, not all of the manipulations involved in these experiments are relevant to the theoretical issues I address in this article. Nevertheless, these experiments can be employed to address the theoretical questions at hand because they incorporate a design frequently employed in studies of polarization effects (i.e., a design in which subjects' independent responses to an issue are recorded, interpersonal communication on the issue is opened, and subjects' final opinions on the issue are recorded), and because they involve three "risky shift" issues that polarization studies also address.

Tetrads experiment. The tetrads experiment involved 50 four-person groups of college students. Group members were asked

to attempt to resolve their initial differences of opinion on various issues. Subjects were randomly assigned to one of five different networks of interpersonal communication. These networks are the star {A–C, B–C, D–C}, kite {A–B, A–C, B–C, C–D}, circle {A–B, B–C, C–D, D–A}, slash {A–B, B–C, C–D, D–A, A–C}, and complete {A–B, B–C, C–D, D–A, A–C, B–D} networks.[4] (See Figure 1.) During the experiment, neither the structure of the communication network nor individuals' positions in the network were altered.

Each group member occupied a private room and was given an issue to consider in isolation from the other three group members. Each person was asked to record an initial opinion on the issue. Group members then discussed their opinions using a simple telephone system. Each subject's telephone displayed the names of persons with whom direct communication was possible. Only dyadic communication was permitted and (depending on the network structure) only certain communication channels could be activated by each subject. For example, in the "slash" network, members B and D could not converse with each other, but each could

---

[4] Along with a chain {A–B–C–D}, these five networks include all the nonisomorphic connected structures (connected graphs) that can occur in a four-person group.

converse with network members A and C. Group members were instructed that they could communicate with other members of the group as frequently as they liked, but that they must communicate at least once with each person whose name was listed on their telephones.

Group members were given up to 20 minutes to discuss the assigned issue. Each group was instructed that attaining consensus was feasible and desirable:

> Your goal is to reach consensus. If it seems difficult to reach consensus, remember that most groups are able to come to some decision if those who disagree will restate their reasons and if the problem is reread carefully.

Upon reaching group consensus or a deadlock, group members were asked to record their final opinions on the issue.

Each group dealt with five issues in sequence. To eliminate crossover effects, the order of the issues was systematically varied among the groups. Three discussion issues were choice dilemmas that appear in the "risky shift" literature: the "Sports," "School," and "Surgery" issues (described below). Two other issues involved making a judgment about an appropriate monetary reward: the "Asbestos" and "Disaster" issues (also described below). Individuals' opinions on these issues were represented by real numbers—subjective probabilities on the "risky shift" issues or dollars on the monetary issues.

Triads experiment. The triads experiment involved 32 three-person groups. Each group dealt with the three "risky shift" issues (Sports, School, and Surgery) in the context of either a complete or a chain communication network.[5] One-half of the triads operated under a "high" pressure condition, and one-half operated under a "low" pressure condition. In the "high" pressure condition, subjects were instructed that attaining consensus was feasible and desirable:

> We would like you to reach an agreement. If at the end of 20 minutes there are remaining differences that you believe might be reconciled, you may have an additional 10 minutes for dis-

cussion. You may terminate the session at any time if you believe that the remaining differences of opinion cannot be reconciled. However, it has been our experience that most discussion groups are able to reach an agreement within the 20- (plus optional 10-) minute time frame.

In the "low" pressure condition, subjects were instructed that any outcome was acceptable:

> When the buzzer sounds a second time it is the signal for you to begin telephone communication with the other person. Now is the time to reconsider your choice. Discuss the situation with the other person. The conversation that you will have may or may not lead you to alter your first opinion, and you may or may not come to an agreement. Any of these outcomes are OK with us. You will have 20 minutes in which to discuss the issue. You may have an additional 10 minutes if you want them.

Dyads experiment. The dyads experiment involved 36 two-person groups. Dyads were given up to 30 minutes to discuss an issue. Eighteen of the dyads were placed under the "high" pressure condition described in the triads experiment, and 18 dyads were placed under the "low" pressure condition. Each dyad dealt with three discussion issues in sequence. Two of these issues were "risky shift" issues—Surgery and School. The third issue, taken from the General Social Survey of the National Opinion Research Center, involved multiple assessments of the level of confidence in various U.S. institutions.

The Discussion Issues

Prior findings on choice shifts and group polarization have been based mainly on choice dilemmas in which subjects assess an acceptable level of risk in adopting a particular course of action. Three of the discussion issues I used involve choice dilemmas that have been used by other investigators and for which there has been some evidence of group polarization. On these three choice dilemmas (Sports, School, and Surgery), subjects were instructed to choose between two alternative courses of action:

> One alternative involves greater risk than the other, while also offering a greater potential reward. Consider the alternatives. Then indicate what probability of success would be nec-

---

[5] In the chain network, the ties are {A–B, B–C} and in the complete network they are {A–B, B–C, A–C}.

essary for you to choose the alternative which is potentially more rewarding, but which also carries a greater degree of risk.

The Sports issue involved a choice between alternative plays in a game between two college teams (one play would tie the game; the other play could win it or lose it). The School issue involved a choice between two Ph.D. programs (one program has a low failure rate; the other program has a higher failure rate but a much better academic reputation). The Surgery issue involved a choice between two courses of medical treatment (one treatment entails little risk but a drastic curtailment of life style; the riskier treatment might bring about a complete cure). The subjects' responses (opinions) were restricted to one of 20 probability values: .05, .10, .15, ..., 1.00. Previous research indicates that subjects have heterogeneous initial opinions on these issues and take them seriously.

Choice shifts and group polarization effects have been observed on issues other than those involving choice dilemmas. I included two issues of monetary reward, involving compensation for asbestos removal and court damages for a chemical spill (the exact wording of these questions will be provided on request). The "Institutions" issue involved a list of 13 institutions and asked for the subject's level of confidence in the people running each of the institutions on a scale of 0 to 100 percent. The institutions listed were: banks and financial institutions, major companies, organized religion, education, the executive branch of the federal government, organized labor, the press, medicine, television, the U.S. Supreme Court, the scientific community, the U.S. Congress, and the military. A subject's attitude on this issue is represented by his or her mean level of confidence for the 13 institutions.

Measure of Relative Interpersonal Influence

The network influence model requires a measure of the relative interpersonal influence of each actor on another. These measures are used in the matrix $\mathbf{C} = [c_{ij}]$ that enters into equation 2. In prior work, I derived a measure of $\mathbf{C}$ from features of the structure of the communication network in which two

actors are situated (Friedkin 1993, 1998). My approach assumed that this structure reflects the power bases of the actors. In this experimental setting, in which communication structures are artificially constrained, a measure of interpersonal influence could be based on a content analysis of the communications that occur among subjects. Here I have adopted a simpler approach that relies on subjects' assessments of the relative influence of other group members on their final opinion on an issue (Hunter 1953; Laumann and Pappi 1976, chap. 5; Merton 1968, chap. 12; Tannenbaum 1974).

This reliance on subjective reports has strengths and weaknesses. One strength is that power bases can translate into direct interpersonal influence only if they are *perceived*. Thus, for example, French and Raven's (1959) exegesis of power bases emphasizes the perceptual mediation of interpersonal influence. Weaknesses involve subjects' bias and random error. For example, actors may underestimate the influence of other actors and inflate their own importance. I control for such bias, however, by using the *relative* values of the reports as a measure of the *relative* influence of other actors. Another concern is whether actors can accurately disentangle the relative influences of multiple actors, especially when there is no dominant actor.[6] To the extent that the subject's judgment is prone to error, the measure of $\mathbf{C}$ will be flawed and errors will be introduced into some predictions.

To solicit information on relative interpersonal influences, subjects were asked (after recording their final opinions on an issue) to estimate the extent to which each group member influenced their final opinion:

> You have been given a total of 20 chips. Each chip represents influence upon your final opinion. Divide the chips into two piles, *Pile A* and *Pile B*. *Pile A* will represent the extent to which

---

[6] An actor may or may not perceive the direct influence of another actor. For instance, an actor may distort or forget the origins of his or her modified opinion on an issue. An actor's opinion may move toward the opinion of another actor who has a negligible total influence on this movement; conversely, his or her opinion may not move toward, and it may even move away from, the opinion of another actor who has had a substantial direct influence.

the conversations you had with the other person influenced your final opinion. *Pile B* will represent the extent to which the conversations you had with the other person did not influence your final opinion. Now consider the extent to which you feel each member of the group influenced the group's final opinion. Divide the chips in *Pile A* into piles for each person according to how much they influenced your final opinion.

The measure of relative interpersonal influence, c_{ij}, is the number of chips that actor i accords to actor j divided by the total number of chips that actor i accorded to all other members of the group (i.e., the number of chips in Pile A).

A Methodological Note on Choice Shifts

Operational definitions of choice shift and group polarization vary. Many studies compare the mean of N individuals' initial opinions with the mean of K groups' decisions (the unanimous or average opinion of a group after discussion) composed of the same N individuals. Some studies compare the mean of N individuals' initial opinions with the mean of the same N individuals' opinions after discussion of the issue. Other studies compare group-level mean opinions before and after discussion. Sometimes different groups are involved in the measures of initial and final opinions in a between-group design. Sometimes a distinction is made between group-level mean opinions after discussion and group-level consensual decisions on the issue after the discussion.

The study of choice shift and group polarization must consider the strong likelihood that the members of a group do not respond independently to a treatment (e.g., group discussion). A test of the difference in the means of N subjects' opinions before and after group discussion does not satisfy the assumptions of independence on which the statistical test for the difference of means depends. This violation of assumptions occurs when the treatment (group discussion) involves subjects who are responding (with revised opinions) to the displayed responses (opinions) of the other members of the group.

A straightforward way to deal with the independence assumption is to use a paired-samples *t*-test in which the absolute values of the gain or loss scores are evaluated at the group level (i.e., a gain or loss is defined as the difference between the mean final opinion and mean initial opinion of the members of a group). When assessing choice shifts, this approach permits a statement about whether *on average* a group's decision (unanimous or average opinion) departs from the group's mean initial opinion.

Special methods are required to detect group polarization, which is a more specific phenomenon than choice shift. A choice shift may occur without any group polarization; moreover, even a positive or negative difference on average between a group's final decision and mean initial opinion can occur without group polarization. For example, a positive difference occurs if the group's final decision (unanimous or average opinion) is higher on the opinion scale than the group's mean initial opinion for each group in the sample. However, a simple escalation or de-escalation of opinions is not group polarization. Group polarization refers to a particular *pattern* of choice shifts: If the mean initial opinion of a group is above (below) a certain point, discussion shifts the average opinion to a more extreme position in the same direction as the group's initial inclination.

The straightforward method of detecting group polarization is to regress a group's mean final opinion on its mean initial opinion. Figure 2 illustrates some possible outcomes of such a regression. For simplicity, I ignore floor and ceiling effects for issues that are scaled between 1 and 0 (for sufficiently high mean initial opinions, group decisions approach 1; for sufficiently low mean initial opinions, group decisions approach 0, regardless of any choice shift and group polarization effects). The baseline in Figure 2 indicates a pattern in which groups converge to the mean of their initial opinions—no choice shift. The other curves—illustrating an escalation, de-escalation, or group polarization of opinions—are generally indicated by an intercept that is significantly different from 0 or by a slope that is significantly different from 1.

I analyze the subjects' behavior on individual issues. Many early studies, especially those employing Stoner's (1961) methodology, document polarization with respect to mean shifts across a set of issues. But it was
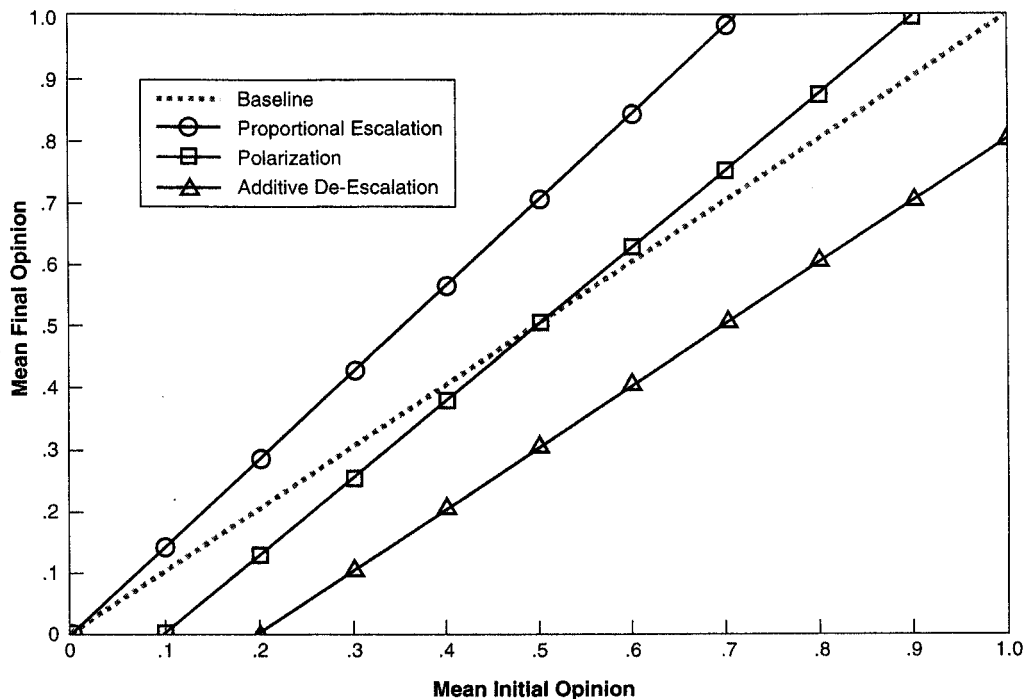
Figure 2. Illustrative Patterns of Choice Shift

found that whether or not groups polarize depends on the particular issue (Cartwright 1971:368). Other studies have documented group polarization with respect to shifts on particular issues (within-issue shifts). Discussions of polarization usually are couched in terms of its implications for the behavior of a *particular* group discussing an issue (Myers 1996:334). I focus on a within-issue analysis because this approach has the most direct bearing on the expected behavior of a small group discussing an issue.

RESULTS

Table 1 presents evidence on the occurrence of choice shifts. First, the groups do not converge to the mean of members' initial opinions. The null hypothesis, that the average absolute value of the choice shifts in a population of groups is 0, is rejected in all cases. Second, group discussion does not reliably shift opinions in one direction or another. A significant difference between the grand mean of subjects' initial opinion and the mean of their groups' collective decision (i.e., the groups' final consensus or mean final opinion) appears only for the Disaster is-

sue in the tetrads.[7] The effect of discussion on the Disaster issue is to substantially lower the monetary reward. Third, the observed group-level mean final opinions, $\bar{y}_{\bullet}^{(\infty)}$, do not differ significantly from the network theory's predicted mean final opinions.

Table 2 reports regression results for two models—Model 1 indicates group polarization and Model 2 tests the adequacy of a social influence network approach to explaining choice shifts. Model 1 is simply

$$\bar{y}_k^{(\infty)} = b_0 + b_1 \bar{y}_k^{(1)} + e_k, \tag{9}$$

where $\bar{y}_k^{(\infty)}$ and $\bar{y}_k^{(1)}$ are, respectively, the mean final opinion and mean initial opinions of group k. Model 1 assesses whether there is a systematic linear pattern in the choice shifts that departs from the baseline assumption of no choice shifts. The null hypothesis is rejected if the intercept is not 0 or the

[7] This type of comparison was common in the "risky shift" literature and was linked to the hypothesis that group decisions are often riskier than individual decisions. Note that the mean of group means is simply the grand mean when the size of the groups is identical.

Table 1. Choice Shifts, by Type of Group and Issue

Type of Group and Issue	Choice Shift $\Delta = \frac{1}{N}\sum_{k=1}^{N} abs\left(\bar{y}_k^{(\infty)} - \bar{y}_k^{(1)}\right)$		Group-Level Mean, Final Opinion $\bar{y}_\bullet^{(\infty)} = \frac{1}{N}\sum_{k=1}^{N} \bar{y}_k^{(\infty)}$				Grand Mean, Initial Opinions $\bar{y}_{\bullet\bullet}^{(1)} = \frac{1}{NG}\sum\sum y_{ik}^{(1)}$	
			Observed		Predicted[c]			
	Δ [a]	S.D.	$\bar{y}_\bullet^{(\infty)}$	S.D.	$\hat{\bar{y}}_\bullet^{(\infty)}$	S.D.	$\bar{y}_{\bullet\bullet}^{(1)}$	S.D.
Tetrads								
Sports	8.60	(7.86)	44.10	(17.54)	46.22	(16.99)	48.70	(23.86)
Asbestos	1.28	(1.67)	9.57	(3.42)	9.48	(3.47)	9.03	(4.46)
School	8.68	(9.26)	45.80	(16.63)	45.35	(16.50)	46.83	(21.68)
Disaster	192.77	(175.03)	145.25[b]	(182.05)	193.62	(183.45)	314.94	(417.07)
Surgery	6.25	(4.55)	69.05	(14.03)	69.06	(13.82	68.10	(23.10)
Triads								
Surgery	5.82	(5.65)	74.16	(14.27)	73.56	(14.76)	71.67	(23.27)
Sports	8.44	(9.46)	47.24	(21.40)	47.69	(21.16)	52.66	(23.55)
School	7.50	(6.13)	54.22	(22.20)	54.44	(21.46)	54.84	(25.26)
Dyads								
School	6.69	(6.95)	71.68	(14.92)	70.31	(15.88)	69.76	(21.83)
Surgery	8.68	(10.53)	55.07	(21.28)	55.52	(17.34)	54.15	(22.83)
Institutions	2.67	(2.63)	70.33	(8.45)	70.85	(7.32)	71.53	(10.32)

Note: Numbers in parentheses are standard deviations. N is the number of groups (50 for tetrads, 32 for triads, and 36 for dyads); G is the number of subjects in each group (4 for tetrads, 3 for triads, and 2 for dyads).

[a] H_0: $\Delta = 0$ is rejected ($p < .05$) in all cases.

[b] H_0: $\bar{y}_\bullet^{(\infty)} - \bar{y}_{\bullet\bullet}^{(1)} = 0$ is rejected ($p < .01$).

[c] H_0: $\bar{y}_\bullet^{(\infty)} - \hat{\bar{y}}_\bullet^{(\infty)} = 0$ is accepted ($p > .05$) in all cases.

slope is not 1. These omnibus tests can detect evidence of group polarization or any systematic departure from the baseline model (i.e., $b_0 = 0$, $b_1 = 1$).

Model 1 indicates that departures from the baseline assumptions are neither ubiquitous nor reliable. For the Disaster issue in tetrads, the confidence limits for the slope (.25, .64) suggests that discussion diminished the average monetary rewards. For the Surgery issue in triads, the confidence limits for the intercept (3.03, 29.05) suggests that discussion tended to increase the average level of assurance demanded for a successful operation, and the limits for the slope (.63, .99) suggests that the more conservative a group is initially, the more likely it is that the group will accept a greater degree of risk. For the Surgery issue in dyads and tetrads, however, the results are consistent with baseline assumptions. For the Sports issue in triads, the confidence limits for the intercept (−29.65, −.11) suggests that discussion tended to increase the average level of risk-taking in a group whereas in tetrads there is no such effect. Finally, for the School issue in dyads, the confidence limits for the intercept (3.28, 31.83) suggests that discussion tended to diminish risk-taking, and the limits for the slope (.58, .98) suggests that the more conservative a group is initially, the more likely it is that the group will accept a higher degree of risk. However, on the School issue in triads and tetrads, the results are consistent with baseline assumptions. In short, although there are choice shifts in these data, the shifts are not patterned in a systematic way among the groups addressing a particular issue.

Model 2 suggests that the network model provides a plausible explanation of how choice shifts are produced. Model 2 regresses the group's mean equilibrium opin-

Table 2. Confidence Limits for Regression Coefficients in Group Polarization and Social Influence Network Models

Type of Group and Issue	Model 1 (Group Polarization) $\bar{y}_k^{(\infty)} = b_0 + b_1 \bar{y}_k^{(1)} + e_k$		Model 2 (Network Influence) $\bar{y}_k^{(\infty)} = b_0 + b_1 \hat{\bar{y}}_k^{(\infty)} + e_k$		R^2		Number of Groups
	Intercept, b_0	Slope, b_1	Intercept, b_0	Slope, b_1	Model 1	Model 2	
Tetrads							
Sports	−16.71, 6.33	.78, 1.24	−5.82, 4.02	.87, 1.07	.624	.889	50
Asbestos	−3.00, 1.51	.90, 1.38	−.20, 1.15	.89, 1.02	.654	.945	50
School	−15.30, 16.56	.63, 1.30	−1.75, 9.44	.81, 1.04	.417	.842	50
Disaster	−69.11, 81.73	.25, .64[b]	−57.40, 26.30	.67, .99[b]	.301	.700	50
Surgery	−10.68, 15.12	.79, 1.17	−3.76, 11.56	.83, 1.05	.697	.864	50
Triads							
Surgery	3.03, 29.05[a]	.63, .99[b]	−.56, 8.63	.89, 1.01	.743	.971	32
Sports	−29.65, −.11[a]	.91, 1.45	−3.47, 2.68	.94, 1.06	.727	.975	32
School	−22.88, .35	.99, 1.40	−4.26, .78	.98, 1.07	.828	.987	32
Dyads							
School	3.28, 31.83[a]	.58, .98[b]	1.94, 22.32[a]	.71, .99[b]	.646	.813	36
Surgery	−13.63, 18.58	.69, 1.25	−15.47, 11.88	.79, 1.26	.586	.697	36
Institutions	−24.47, 1.05	.97, 1.32	−16.38, 2.21	.96, 1.22	.835	.895	36

[a] Does not include 0.

[b] Does not include 1.

ion on the predicted mean opinion for the group

$$\bar{y}_k^{(\infty)} = b_0 + b_1 \hat{\bar{y}}_k^{(\infty)} + e_k, \qquad (10)$$

where $\bar{y}_k^{(\infty)}$ and $\hat{\bar{y}}_k^{(\infty)}$ are, respectively, the observed and predicted mean final opinions of group k. Compared with Model 1, Model 2 predicts mean opinions with a higher degree of accuracy. Model 2 explains an additional 11.1 percent to 42.5 percent of the variance among groups, and the 95-percent confidence limits are consistent with parameter values of 0 for the intercepts and of 1 for the slopes. Exceptions occur for the Disaster issue in tetrads where the 95-percent confidence limits for the slope is (.67, .99), and for the School issue in dyads where the 95-percent confidence limits for the intercept and slope are (1.94, 22.32) and (.71, .99) respectively. In these two cases, there may be some effects on opinions that cannot be attributed to the network of interpersonal influences. These are marginal exceptions that, for the most part, disappear under a more conservative test (Bonferroni joint confidence limits for the slope and intercept). Under this test only one result is inconsistent with the assertion that a group's social influence network explains escalation, de-escalation, and group polarization on issues (i.e., an intercept significantly different from 0 in dyads on the School issue).

A potential weakness of the measure of the group's influence network is that it is based on information subjects provide *after* they have completed their discussion of an issue. The influence network formed on one issue may carry over into the group's discussion of another issue, however, especially if the two issues do not involve substantially different types of expertise and interests. Table 3 reports the findings (comparable to Model 2 in Table 2) from regressing the group's mean equilibrium opinion on the predicted mean opinion for the group, where the predicted mean final opinion is based on the influence network for group k for the immediately preceding issue in the experimental trials. For instance, in the tetrads experiment the five issues (Sports, Asbestos, School, Disaster, and Surgery) were rotated so that each issue was evenly distributed in the sequence of issue discussions. Hence, on each issue there are 40 groups for which there is an im-

Table 3. Confidence Limits for Regression Coefficients in Model 2 Based on the Influence Network from Groups' Previous Experimental Trial

Type of Group and Issue	95-Percent Confidence Limits for:		R_2	Number of Groups
	Intercept, b_0	Slope, b_1		
Tetrads				
Sports	–5.94, 9.83	.72, 1.02	.783	40
Asbestos	1.15, 3.99[a]	.62, .90[b]	.761	40
School	–1.50, 11.82	.78, 1.06	.823	40
Disaster	–70.59, 28.21	.56, .93[b]	.644	40
Surgery	–4.03, 13.26	.80, 1.05	.858	40
Triads				
Surgery	.05, 20.55[a]	.73, 1.01	.927	16
Sports	–12.00, 7.87	.79, 1.14	.910	16
School	–10.57, 4.91	.90, 1.17	.895	32
Dyads				
School	1.11, 50.40[a]	.31, .96[b]	.426	24
Surgery	–2.41, 47.09	.21, 1.06	.305	24
Institutions	–36.20, 11.32	.84, 1.51	.708	24

Note: The School issue was not rotated for triads.

[a] Does not include 0.

[b] Does not include 1.

mediately prior discussion of a different issue, and the influence network for that immediately prior discussion was used to predict the outcomes of the subsequent discussion. The results show that the influence network model significantly predicts the mean opinion of the members of a group on each of the issues that were dealt with by the tetrads, triads, and dyads.

DISCUSSION

The present findings support three conclusions. First, choice shifts are a ubiquitous product of interpersonal interactions on issues; however, group polarization (the tendency for such shifts to enhance the group's initial position) is neither ubiquitous nor reliable.[8] Second, a choice shift can be viewed

_____

[8] Group polarization may be characteristic of "open" discussion groups in which communication is not restricted to dyadic interactions. Such groups were not investigated in the present study. Even if there were a reliable occurrence of group polarization in such groups, such an occurrence would threaten the present approach only if it could not be explained by the influence structures of these groups.

as a byproduct of the group's influence structure and an outcome of the same process that sometimes produces convergence on the mean of the group's initial opinions. Third, the idea that pressures toward uniformity cause a group's opinions to converge to the mean of members' initial opinions should be replaced with a more general idea that such pressures cause a group's opinions to converge to a weighted average of members' initial opinions.

Cartwright's (1971) analysis of the "risky shift" literature not only suggested that there was no reliable "risky shift" but also suggested a reconceptualization of "risky shifts." He suggested that the question was not "Why does group discussion produce riskier decisions?" but rather "Why does group discussion produce choice shifts?" Although the field has moved to a broader conceptualization in which choice shifts are now the focus of study, research is still wedded to a paradigm in which some group-level condition (a norm, a decision rule, a pool of arguments, a distribution of initial opinions) determines group outcomes. My study suggests that choice shifts should be viewed as an aggregate manifestation of an

influence process, and that the focus should not be on the choice shift itself (which derives from the influence process) but on the opinion changes that occur for each individual in the group. From this perspective, the crucial phenomenon is the network of interpersonal influence and the process by which these interpersonal influences modify actors' opinions. These findings support Cartwright's (1971) proposal that the formal theory of social power developed by French (1956) and Harary (1959) offers a promising starting point.

Social influence network theory is consistent with social decision scheme theory because, from a purely formal point of view, it asserts that group members form revised opinions that are weighted averages of the group's initial opinions. This theory demonstrates the feasibility of incorporating individual differences into social decision scheme theory.

This network approach is also consistent with social comparison theory. An interpersonal influence network may be based on the visibility (mere awareness) of the distribution of opinions in a group, and a choice shift may result when group members accord different weights to the alternative opinions of which they are aware. Festinger (1954) described a set of hypotheses about the relationship between interpersonal influence and opinion discrepancies. Subsequent studies from various theoretical perspectives touch on his thesis that actors' locations in the group's distribution of initial opinions substantially shape the group's network of interpersonal influences. I advance Festinger's thesis because his results point to the influence network—**W**—as a key theoretical construct.

Clearly, a persuasive-arguments perspective on choice shifts is also consistent with an influence network approach. I suggest that the effects of persuasive arguments can be conceptualized at a dyadic level as a pattern of interpersonal influences. Knowing who is persuading whom should improve an account of the direction of a group's opinion changes. The cognitive processing perspective could be broadened to include a wider domain of effects on interpersonal influence. For example, individuals may shift their opinions toward the opinion of a person who has simply displayed an opinion if that person has power bases (French and Raven 1959) or status characteristics (Berger, Conner, and Fisek 1974) that make that displayed opinion especially salient. Hence, competition between researchers in the social-comparison and persuasive-arguments camps could continue in the context of inquiries concerned with the thesis that the distribution of initial opinions in a group is an important foundation of a group's influence network.

Finally, my study is consistent with self-categorization theory. The existence of a normative or prototypical position should be manifested in the group's pattern of interpersonal influences. For example, proximity to a normative or prototypical position may be associated negatively with susceptibility to interpersonal influence (an actor who is proximate to such a position may be less susceptible to interpersonal influence than an actor who is distant from the position) and associated positively with interpersonal influence (the opinion of an actor who is proximate to such a position may be accorded more weight than the opinion of an actor who is distant from the position). If a normative position on an issue exists prior to group discussion, then it should be reflected in the group's initial opinion distribution. For example, generally accepted norms should produce consensus or near consensus in the distribution of a group's initial opinions. The present work advances self-categorization theory by specifying three possible causal routes for normative influences on opinions—via an effect on persons' initial opinions, susceptibilities to interpersonal influence, and weights accorded to opinions of particular other members of a group.

My study reasserts an older tradition in which the structure of a group's influence network is viewed as an important construct in the study of group dynamics. A model of the social influence process must allow for inequalities of interpersonal influence, for situations in which there is not a massive initial consensus, for situations in which no actor's opinion is fixed, and for situations in which the influence process does not necessarily produce a consensus. Group polarization appears to be an unreliable phenomenon because it depends on the groups' particular

structures of interpersonal influence. Choice shifts *are* a reliable phenomenon because influence structures tend to involve inequalities of interpersonal influence, which imply such shifts (Ridgeway and Walker 1995). Conditions that affect group members' susceptibilities to interpersonal influence (group cohesion, attitude strength) may affect the magnitude of a choice shift, and conditions that affect the relative influence of the members of a group (members' power bases and status characteristics) may affect the direction of a choice shift. Hence, this theory provides a promising integrative framework that is in line with classical studies of influence and conformity, for assessing hypotheses concerned with the etiology of opinion changes and group outcomes.

Analysis of the social network structure of groups became prominent in sociology while the "cognitive revolution" reinforced a neglect of group dynamics in psychology. Bringing a network perspective to bear on neglected topics in group dynamics may help to reinvigorate this field of inquiry. For modern structural social psychology (Lawler et al. 1993), this development would entail enlarging the study of power to encompass influence networks that cannot easily be fitted into the social exchange paradigm and extending the study of status structures into contexts in which status is a relational (rather than an individual) attribute and is founded on various bases, including but not restricted to actors' performance expectations. Because networks of interpersonal influence, broadly construed, enter into many domains of sociological inquiry, theories concerned with the contributions of these networks have widespread implications. As I have demonstrated, a network perspective does not represent an attack on existing work that overturns and disrupts extant findings, but instead envelops and enhances extant findings by attending to the group structure within which interpersonal processes unfold. The pursuit of a more refined specification of structural effects, which involves an elaborated theoretical understanding of how interpersonal processes unfold in complex social structures, is a goal in many fields of sociological inquiry and supports the enduring interest of sociologists in social network structure and process.

Noah E. Friedkin is Professor of Sociology at the University of California, Santa Barbara. His current research focuses on the development and application of social influence network theory. He also continues to pursue research on social exchange networks. His most recent works include: A Structural Theory of Social Influence *(New York: Cambridge University Press, 1998),* "Social Influence Networks and Opinion Change" *(with Eugene Johnsen,* Advances in Group Processes, *1999, vol. 16, pp. 1–29) and* "Unequally Valued Exchange Relations" *(with Phillip Bonacich,* Social Psychology Quarterly, *1998, vol. 61, pp. 160–71).*

Appendix A. Estimating Actors' Susceptibilities to Influence

The various approaches for operationalizing the social influence theory described by equation 1 depend on the availability of measures for the theoretical constructs. Friedkin and Johnsen (1999) have developed the following approach for estimating actors' susceptibilities to interpersonal influence (**A**) when data are available on actors' initial opinions, final opinions, and relative interpersonal influences (i.e., $\mathbf{y}^{(1)}$, $\mathbf{y}^{(\infty)}$, and **C** respectively).

Assuming equilibrium, the scalar equation of the reduced-form, equation 5, is

$$y_i^{(\infty)} = a_{ii}(1 - a_{ii})y_i^{(\infty)} + a_{ii}^2 \sum_{j \neq i} c_{ij} y_j^{(\infty)}$$
$$+ (1 - a_{ii})y_i^{(1)}, \qquad (\text{A-1})$$

from which it follows that

$$y_i^{(\infty)} - y_i^{(1)} = a_{ii}(y_i^{(\infty)} - y_i^{(1)}) + a_{ii}^2(\overline{y}_i^{(\infty)} - y_i^{(\infty)}), \quad (\text{A-2})$$

and

$$(1 - a_{ii})(y_i^{(\infty)} - y_i^{(1)}) = a_{ii}^2(\overline{y}_i^{(\infty)} - y_i^{(\infty)}), \qquad (\text{A-3})$$

where

$$\overline{y}_i^{(\infty)} = \sum_{j \neq i} c_{ij} y_j^{(\infty)} \qquad (\text{A-4})$$

is a weighted average of the others' settled opinions. Hence, for $\overline{y}_i^{(\infty)} - y_i^{(\infty)} \neq 0$,

$$\frac{a_{ii}^2}{1 - a_{ii}} = \frac{y_i^{(\infty)} - y_i^{(1)}}{\overline{y}_i^{(\infty)} - y_i^{(\infty)}} \equiv \Delta_i, \qquad (\text{A-5})$$

and

$$a_{ii} = \frac{-\Delta_i \pm \sqrt{\Delta_i^2 + 4\Delta_i}}{2}. \qquad (\text{A-6})$$

The a_{ii} computed from equation A-6 is a complex number for $-4 < \Delta_i < 0$, greater than 1 for $\Delta_i \leq -4$, and less than 1 for $0 \leq \Delta_i$. See Appendix Table A.

But assumptions $w_{ii} = 1 - a_{ii}$ and $0 \leq w_{ii} \leq 1$ imply that $0 \leq a_{ii} \leq 1$. Hence, for each real-valued Δ_i, the estimate of a_{ii} is selected to be the real number in the legitimate range [0,1] that is numerically the

Appendix Table A. Solutions for Equation A-6

Value of Δ_i	$a_{ii} = \dfrac{-\Delta_i + \sqrt{\Delta_i^2 + 4\Delta_i}}{2}$	$a_{ii} = \dfrac{-\Delta_i - \sqrt{\Delta_i^2 + 4\Delta_i}}{2}$
$\Delta_i \le -4$	$a_{ii} \ge 2$	$1 < a_{ii} \le 2$
$-4 < \Delta_i < 0$	Complex number	Complex number
$\Delta_i \ge 0$	$a_{ii} < 1$	$a_{ii} \le 0$

closer (in the complex plane) to either of the a_{ii} computed from equation A-6. The resulting solution set for $\bar{y}_i^{(\infty)} - y_i^{(\infty)} \ne 0$ is:

(a) $\Delta_i \le -2 \Rightarrow a_{ii} = 1$,

(b) $-2 < \Delta_i < 0 \Rightarrow a_{ii} = \dfrac{-\Delta_i}{2}$,

(c) $\Delta_i \ge 0 \Rightarrow a_{ii} = \dfrac{-\Delta_i + \sqrt{\Delta_i^2 + 4\Delta_i}}{2}$.

Note that this solution assumes $\bar{y}_i^{(\infty)} - y_i^{(\infty)} \ne 0$, an assumption that will be violated when there is an equilibrium consensus or, in the absence of such consensus, when actor i has an equilibrium opinion that is the weighted average, $\bar{y}_i^{(\infty)}$, of the others' equilibrium opinions.

When $\bar{y}_i^{(\infty)} - y_i^{(\infty)} = 0$ and $y_i^{(\infty)} - y_i^{(1)} \ne 0$, then $a_{ii} = 1$, an implication that follows from equation A-3. When $\bar{y}_i^{(\infty)} - y_i^{(\infty)} = 0$ and $y_i^{(\infty)} - y_i^{(1)} = 0$, then a_{ii} can be any value. Actor i's opinion has not changed either because actor i was not susceptible to interpersonal influence or because he or she was susceptible to such influence but remained in the same position as a result of balancing cross-pressures. The former situation (which corresponds to $a_{ii} = 0$) is more likely than the latter. Therefore, it is assumed that $a_{ii} = 0$ with the understanding that this assumption is a potential source of error in the model.

REFERENCES

Allport, Floyd H. 1924. *Social Psychology*. New York: Houghton Mifflin.

Anderson, Norman H. 1996. *A Functional Theory of Cognition*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, Norman H. and Cheryl C. Graesser. 1976. "An Information Integration Analysis of Attitude Change in Group Discussion." *Journal of Personality and Social Psychology* 34: 210–22.

Asch, Solomon E. 1951. "Effects of Group Pressure upon the Modification and Distortion of Judgement." Pp. 117–90 in *Groups, Leadership and Men*, edited by M. H. Guetzkow. Pittsburgh, PA: Carnegie Press.

Baron, Robert S., Sieg I. Hoppe, Chaun Feng Kao, Bethany Brunsman, Barbara Linneweh, and Diane Rogers. 1996. "Social Corroboration and Opinion Extremity." *Journal of Experimental Social Psychology* 32:537–60.

Berger, Joseph, Thomas L. Conner, and M. Hamit Fisek. 1974. "Expectation States Theory: A Theoretical Research Program." Cambridge, MA: Winthrop.

Berger, Joseph, M. Hamit Fisek, Robert Z. Norman, and Morris Zelditch Jr. 1977. *Status Characteristics and Social Interaction*. New York: Elsevier.

Berger, Joseph, Cecilia L. Ridgeway, M. Hamit Fisek, and Robert Z. Norman. 1998. "The Legitimation and Delegitimation of Power and Prestige Orders." *American Sociological Review* 63:379–405.

Burnstein, Eugene. 1982. "Persuasion as Argument Processing." Pp. 103–24 in *Contemporary Problems in Group Decision-Making*, edited by J. H. Davis and G. Stocher-Kreichgauer. New York: Academic Press.

Burnstein, Eugene and Amiram Vinokur. 1973. "Testing Two Classes of Theories about Group-Induced Shifts in Individual Choice." *Journal of Experimental Social Psychology* 9: 123–37.

———. 1977. "Persuasive Argumentation and Social Comparison as Determinants of Attitude Polarization." *Journal of Experimental Social Psychology* 13:315–32.

Burt, Ronald S. 1982. *Toward a Structural Theory of Action*. New York: Academic Press.

Cartwright, Dorwin. 1971. "Risk Taking by Individuals and Groups: An Assessment of Research Employing Choice Dilemmas." *Journal of Personality and Social Psychology* 20:361–78.

———. 1973. "Determinants of Scientific Progress: The Case of Research on the Risky Shift." *American Psychologist* 22:222–31.

Cartwright, Dorwin and Alvin Zander. 1968. "Group Dynamics." New York: Harper and Row.

Clark, Russell D. 1971. "Group-Induced Shift Toward Risk." *Psychological Bulletin* 76:251–70.

Cook, Karen S., Gary Alan Fine, and James S. House, eds. 1995. *Sociological Perspectives on Social Psychology*. Boston, MA: Allyn and Ba-

con.

Davis, James H. 1973. "Group Decision and Social Interaction: A Theory of Social Decision Schemes." *Psychological Review* 80:97–125.

DeGroot, Morris H. 1974. "Reaching a Consensus." *Journal of the American Statistical Association* 69:118–21.

Dion, Kenneth L., Robert S. Baron, and Norman Miller. 1970. "Why Do Groups Make Riskier Decisions than Individuals?" Pp. 305–77 in *Advances in Experimental Social Psychology*, edited by L. Berkowitz. San Diego, CA: Academic.

Doreian, Patrick. 1981. "Estimating Linear Models with Spatially Distributed Data." Pp. 359–88 in *Sociological Methodology*, edited by S. Leinhardt. San Fransico, CA: Jossey-Bass.

Erbring, Lutz and Alice A. Young. 1979. "Individuals and Social Structure: Contextual Effects as Endogenous Feedback." *Sociological Methods and Research* 7:396–430.

Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations* 7:117–40.

Festinger, Leon, Stanley Schachter, and Kurt Back. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. New York: Harper.

French, John R. P., Jr. 1956. "A Formal Theory of Social Power." *The Psychological Review* 63:181–94.

French, John R. P., Jr. and Bertram Raven. 1959. "The Bases of Social Power." Pp. 150–67 in *Studies of Social Power*, edited by D. Cartwright. Ann Arbor, MI: Institute for Social Research.

Friedkin, Noah E. 1990. "Social Networks in Structural Equation Models." *Social Psychology Quarterly* 53:316–28.

———. 1991. "Theoretical Foundations for Centrality Measures." *American Journal of Sociology* 96:1478–1504.

———. 1993. "Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study." *American Sociological Review* 58:861–72.

———. 1998. *A Structural Theory of Social Influence*. New York: Cambridge University Press.

Friedkin, Noah E. and Eugene C. Johnsen. 1990. "Social Influence and Opinions." *Journal of Mathematical Sociology* 15:193–206.

———. 1997. "Social Positions in Influence Networks." *Social Networks* 19:209–22.

———. 1999. "Social Influence Networks and Opinion Change." *Advances in Group Processes* 16:1–29.

Goethals, George R. and Mark P. Zanna. 1979. "The Role of Social Comparison in Choice Shifts." *Journal of Personality and Social Psychology* 37:1469–76.

Graesser, Cheryl C. 1991. "A Social Averaging Theorem for Group Decision Making." Pp. 1–40 in *Contributiions to Information Integration Theory*, vol. 2, edited by N. H. Anderson. Hillsdale, NJ: Lawrence Erlbaum.

Harary, Frank. 1959. "A Criterion for Unanimity in French's Theory of Social Power." Pp. 168–82 in *Studies in Social Power*, edited by D. Cartwright. Ann Arbor, MI: Institute for Social Research.

Hunter, Floyd. 1953. *Community Power Structure*. Durham, NC: University of North Carolina.

Isenberg, Daniel J. 1986. "Group Polarization: A Critical Review and Meta-Analysis." *Journal of Personality and Social Psychology* 50:1141–51.

Kerr, Norbert L. 1981. "Social Transition Schemes: Charting the Group's Road to Agreement." *Journal of Personality and Social Psychology* 41:684–702.

———. 1992. "Group Decision Making at a Multialternative Task: Extremity, Interfaction Distance, Pluralities, and Issue Importance." *Organizational Behavior and Human Decision Processes* 52:64–95.

Lamm, Helmut and David G. Myers. 1978. "Group-Induced Polarization of Attitudes and Behavior." *Advances in Experimental Social Psychology* 11:145–95.

Latane, Bibb. 1981. "The Psychology of Social Impact." *American Psychologist* 36:343–56.

Laughlin, Patrick R. 1980. "Social Combination Processes of Cooperative Problem-Solving Groups on Verbal Intellective Tasks." Pp. 127–55 in *Progress in Social Psychology*, edited by M. Fishbein. Hillsdale, NJ: Erlbaum.

Laumann, Edward O. and Franz U. Pappi. 1976. *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: Academic Press.

Lawler, Edward J., Cecilia Ridgeway, and Barry Markovsky. 1993. "Structural Social Psychology and the Micro-macro Problem." *Sociological Theory* 11:268–90.

Mackie, Diane M. 1986. "Social Identification Effects in Group Polarization." *Journal of Personality and Social Psychology* 50:720–28.

Marsden, Peter V. and Noah E. Friedkin. 1994. "Network Studies of Social Influence." Pp. 3–25 in *Advances in Social Network Analysis*, vol. 22, edited by S. Wasserman and J. Galaskiewicz. Thousand Oaks, CA: Sage.

McGarty, Craig, John C. Turner, Michael A. Hogg, Barbara David, and Margaret S. Wetherell. 1992. "Group Polarization as Conformity to the Prototypical Group Member." *Bristish Journal of Social Psychology* 31:1–20.

Merton, Robert K. 1968. *Social Theory and So-

cial Structure. New York: Free Press.

Myers, David G. 1996. *Social Psychology*. New York: McGraw-Hill.

Newcomb, Theodore M. 1961. *The Acqaintance Process*. New York: Holt, Rinehart and Winston.

Pruitt, Dean G. 1971. "Choice Shifts in Group Discussion: An Introductory Review." *Journal of Personality and Social Psychology* 20:339–60.

Ridgeway, Cecilia L., Elizabeth Heger Boyle, Kathy J. Kuipers, and Dawn T. Robinson. 1998. "How Do Status Beliefs Develop? The Role of Resources and Interactional Experience." *American Sociological Review* 63:331–50.

Ridgeway, Cecilia L. and Henry A. Walker. 1995. "Status Structures." Pp. 281–310 in *Sociological Perspectives on Social Psycholgy*, edited by K. S. Cook, G. A. Fine, and J. S. House. Boston, MA: Allyn and Bacon.

Sanders, Glenn S. and Robert S. Baron. 1977. "Is Social Comparision Irrelevant for Producing Choice Shifts?" *Journal of Experimental Social Psychology* 13:303–14.

Sherif, Muzafer. 1936. *The Psychology of Social Norms*. New York: Harper.

Stasser, Garold, Norbert L. Kerr, and James H. Davis. 1989. "Influence Processes and Consensus Models in Decision-Making Groups." Pp. 279–326 in *Psychology of Group Influence*, ed-

ited by P. B. Paulus. Hillsdale, NJ: Lawrence Erlbaum.

Stoner, James A. 1961. "A Comparison of Individual and Group Decisions Involving Risk." M.A. Thesis, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

Tannenbaum, Arnold S. 1974. *Hierarchy in Organizations*. San Francisco, CA: Jossey-Bass.

Turner, John C. 1985. "Social Categorization and the Self-Concept: A Social Cognitive Theory of Group Behavior." Pp. 77–122 in *Advances in Group Processes*, vol. 2, edited by E. J. Lawler. Greenwich, CT: JAI Press.

———. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, England: Basil Blackwell.

Turner, John C. and Penelope J. Oakes. 1989. "Self-Categorization Theory and Social Influence." Pp. 233–75 in *Psychology of Group Influence*, edited by P. B. Paulus. Hillsdale, NJ: Lawrence Erlbaum.

Vinokur, Amiram. 1971. "Review and Theoretical Analysis of the Effects of Group Processes upon Individual and Group Decisions Involving Risk." *Psychological Bulletin* 76:231–50.

Zuber, Johannes, Helmut Crott, and Joachim Werner. 1992. "Choice Shift and Group Polarization: An Analysis of the Status of Arguments and Social Decision Schemes." *Journal of Personality and Social Psychology* 62:50–61.