Computational & Mathematical Organization Theory 5:2 (1999): 129–145

Structural Learning: Attraction and Conformity in Task-Oriented Groups

JAMES A. KITTS AND MICHAEL W. MACY
Department of Sociology, Cornell University, 323 Uris Hall, Ithaca, NY 14853
email: jak20@cornell.edu

ANDREAS FLACHE
Department of Philosophy, University of Bayreuth, Germany

Abstract

This study extends previous research that showed how informal social sanctions can backfire when members prefer friendship over enforcement of group norms. We use a type of neural network to model the coordination of informal social control in a small group of adaptive agents confronted with a social dilemma. This model incorporates two mechanisms of social influence, informal sanctions and imitation. Both mechanisms vary with the strength of the social tie between source and target. Previous research focused on the effects of social sanctions. Here, we demonstrate a curvilinear effect of imitation on compliance with prosocial norms. Moderate doses of imitation reduce the coordination complexity of self-organized collective action and help the network achieve satisfactory levels of cooperation. High doses, however, undermine the agent-based learning required to find cooperative solutions. Increasing group size also diminishes compliance due to increased complexity, with larger groups requiring more imitation to overcome the coordination problem.

Keywords: collective action, informal control, social influence, neural networks, computer simulation

1. Introduction

Expected utility theory predicts that an individual member of a group will contribute to the common good only if she receives a *quid pro quo,* or a commensurate payoff that is contingent upon her contribution to the group. In the case of a public good, or a non-excludable group reward, we should expect individuals to shirk and thus “free ride” (Olson, 1965) on others’ efforts. The members then find themselves trapped in a “social dilemma” (Dawes, 1980), in which individuals’ pursuit of self-interest aggregates into collectively suboptimal outcomes.

The need for private incentives suggests that groups may circumvent social dilemmas through informal sanctions, such as social approval. Indeed, social psychologists (e.g., Homans, 1974) argue that members of a production team may actually work *harder* under a group-based reward scheme, as members will serve as taskmasters, praising hard workers and ostracizing shirkers. Conversely, when a group uses a piece-rate scheme to reward workers individually for their efforts, informal norms may emerge to prevent “rate-busters” from pursuing personal gain at the expense of the other workers.

In extending this theory, most researchers have assumed that this effect of social pressure on collective action is moderated by members' dependence on the group for social approval (Seashore, 1954; Festinger et al., 1950; for an overview of empirical studies, see Lott and Lott, 1965). That is, as member dependence increases, social sanctions will become more effective as catalysts of collective action.

However, Flache and Macy (1996) have argued that this theory of dependent compliance is flawed. The potential collective benefit from effective social control is no guarantee that members of a group-rewarded production team will be willing to enforce work norms, due to the "second-order free-rider problem" (Oliver, 1980). Friends with few outside sources of social support may be reluctant to risk personal relationships by using approval as an instrument of social control. Flache and Macy call this reluctance "the weakness of strong ties."

Computer models of the evolution of friendship networks support this conclusion (Flache, 1996; Flache and Macy, 1996; Macy et al., 1997).[1] These simulation studies confirm Homans' prediction that social pressures may solve social dilemmas if approval is used to selectively reward good behavior. However, if agents are allowed the option to offer unconditional approval to their peers, dyads tend to form in which both sides approve of each other while neither works for the group good. The greater coordination complexity of groupwise cooperation may cause social pressures to flow into the maintenance of personal relationships, at the expense of team productivity. Thus, social pressure can "backfire," undermining compliance.

Our work extends this line of theoretical research by relaxing two assumptions. First, we allow that members of a group-rewarded work team may not all have equal regard for one another; instead, social ties may vary in strength. Second, we allow social influence to occur through imitation as well as through the application of sanctions. We now elaborate each of these innovations.

1.1. Variation in the Strength of Ties

Unlike Flache and Macy (1996), who assume relational homogeneity, we allow the strength of social ties to vary over time. We may think of tie strength as a measure of social distance or probability of interaction (Carley, 1991; McPherson and Smith-Lovin, 1987). Based on the *homophily* principle, that "likes attract" (Homans, 1951; Blau, 1977; Blau and Schwartz, 1984), we assume that individuals will grow closer to those with similar behaviors, compared to those with whom they tend to disagree. Specifically, we assume that agents grow closer (or are "attracted") to those who agree with them about whether to work or shirk and about whom to approve or reprimand.

Dyadic attraction, in turn, determines the potency of social sanctions exchanged. We assume that a sanction received through a strong tie is subjectively more potent than an identical sanction received through a weak tie. With this possibility of varying attraction (or social proximity), deviants gain the ability to "avoid" those who disapprove of their behavior, rather than complying with group norms. Moreover, they may build friendships (or strong ties) with those who approve of their deviant behavior. This pattern of deviant solidarity has been widely reported in studies of "counter-cultures" (Yinger, 1982; Willis, 1978; Macleod, 1995).

1.2. Imitation of Role Models

In an early study, Macy (1990) showed that groups of adaptive agents may learn to cooperate in social dilemmas by "random walk" into a critical mass (the number who must cooperate to make the outcome self-reinforcing). He then showed (Macy, 1991) that the coordination complexity of groupwise cooperation is minimized in large groups when agents imitate others (creating a "threshold effect"). This suggests that the results reported by Flache and Macy need to be tested under conditions in which imitation facilitates groupwise coordination of solutions.

In this study, we therefore assume that agents also influence others by modeling behaviors. This means that learning is social, a mixture of influences based on individual experience (reinforcement) and imitation (or what we call "mimetic influence"). As with social sanctions, we model mimetic influence as conditioned by the strength of the social tie between source and target. Rather than assuming that imitation is indiscriminate (Macy, 1991), we assume that individuals are more likely to imitate "friends" than "strangers," as represented by the strength of their relation. Friends thus constitute a reference group for normative guidance. Of course, if the reference group does not agree on a given behavior, then imitation will be difficult or even arbitrary. For this reason, mimetic influence on a given agent i requires a substantial agreement in i's reference group. This agreement corresponds to what Turner and Killian (1957) call an "emergent norm."

The combination of homophily and imitation creates a self-reinforcing dynamic in which similarity increases the probability of interaction and interaction reduces differences between the interactants. Carley (1991) and Axelrod (1997) have used computer simulation to study this dynamic and its effects on cultural diffusion. Carley assumes that shared knowledge promotes exchange of information, which in turn increases knowledge overlap. Axelrod models an array of agents that are most likely to interact with (and thus copy cultural traits of) culturally similar neighbors. Our work takes this simple dynamic a step further by allowing agents to learn from experience, by considering influence through social sanctions as well as imitation, and by introducing tension between individual and collective interests in the diffusion of a trait.

2. Instrumental, Affective and Normative Dependence

We model members of a small, group-rewarded production team who must repeatedly choose whether to contribute to productivity in order to earn a "group wage" or to "free ride" on the efforts of others. We assume individuals may differ from one another in their interest in this public good. The stronger this interest, the more dependent the agent on cooperation by other members of the group. We call this "instrumental dependence."

While each individual agent will always be better off shirking and letting other members produce the group wage, instrumental dependence gives members a "regulatory interest" (Heckathorn, 1988) in the behavior of others that may lead them to sanction one another to encourage compliance with a formal production norm. Following Flache and Macy (1996), we assume that the value of such informal sanctions is moderated by agents' dependence on others in the group for social support, or "affective dependence." We formalize these two concepts in Eqs. (6) and (7) below, where we represent instrumental dependence as parameter γ and affective dependence as parameter β.

Mimetic influence is conditioned by an agent's "normative dependence," or the extent to which an agent relies on its reference group to provide role models for behavior. At higher levels of normative dependence, agents rely less on the lessons from their own experience (including sanctions) and more on imitating their friends. Normative dependence then provides the fulcrum for weighing individual learning against social observation in choosing among alternative courses of action. This is formalized as parameter δ in Eq. (3) below.

This study proposes a type of neural network model to extend the theoretical inquiry begun by Flache and Macy (1996). While they varied affective dependence (relative to instrumental dependence) among groups of equal size and relational strength, we hold both affective dependence and instrumental dependence constant at the group level and manipulate normative dependence in dynamic networks of varying sizes. We also allow social ties to vary in strength over time.

3. An Attractor Neural Network Model

Our computational model incorporates an attractor neural network design (Churchland and Sejnowski, 1994; Quinlan, 1991), originally developed and investigated by Hopfield (1982, 1984; Hopfield and Tank, 1985).[2] This class of models generally includes fully connected networks, in which each node is linked to every other node through ties with endogenous "weights" corresponding to the strength of each tie. Like other neural networks, attractor networks "learn" by adjusting these weights over time, in the absence of global coordination. Through a process of Hebbian learning (Hebb, 1949), each weight w_{ij} adjusts over time as a function of the correlation of the states of nodes i and j. Specifically, Hebbian learning implies the following rules:

- To the extent that nodes i and j adopt the same states at the same time, the weight of their common tie will increase until it approaches some upper limit (e.g., 1.0).
- To the extent that nodes i and j simultaneously adopt different states, the weight of their common tie will decrease until it approaches some lower limit (e.g., 0.0).

Although Hebbian learning was developed to study memory in cognitive systems, it corresponds to the homophily principle in social psychology (Homans, 1951) and social network theory (McPherson and Smith-Lovin, 1987), which holds that agents tend to be attracted to those whom they more closely resemble. This hypothesis is also consistent with Structural Balance Theory (Cartwright and Harary, 1966; Heider, 1958) and has been widely supported in studies of interpersonal attraction and interaction, where it has been called "the Law of Attraction" (Byrne, 1971; Byrne and Griffit, 1966).[3]

Unlike common "feed-forward" neural networks (Rummelhart and McClelland, 1988), which are organized into hierarchies of *input*, *hidden*, and *output* nodes, attractor (or "feed-lateral") networks are internally undifferentiated. Nodes differ only in their states and in their relational alignments, but they are functionally identical.

Without input units to receive directed feedback from the environment, these models are "unsupervised" and thus have no centralized mechanism to coordinate learning of efficient solutions. In the absence of formal training, each node operates using a set of behavioral rules or functions that compute changes of state ("decisions") in light of available information. Zeggelink (1994) calls these "object-oriented models," where each agent receives input from other agents and may transform these inputs into a change of state, which in turn serves as input for other agents.

An important agent-level rule that characterizes attractor networks is that individual nodes seek to minimize "energy" (also "stress" or "dissonance") across all relations with other nodes. This adaptation occurs in discrete stages. In the *action* phase, nodes change their states to maximize similarity with nodes to which they are strongly connected. In the *learning* phase, they update their weights to strengthen ties to similar nodes. Thus, beginning with some (perhaps random) configuration, the network proceeds to search over an optimization landscape as nodes repeatedly cycle through these changes of weights and states.

Ultimately, these systems are able to locate stable configurations (called "attractors"), for which any change of state or weight would result in a net increase in stress for the affected nodes. Hopfield (1982) compares these equilibria to memories, and shows that these systems of undifferentiated nodes can learn to implement higher-order cognitive functions. However, although the system may converge at a stable equilibrium that allows all nodes to be locally satisfied (i.e., a "local optimum"), this does not guarantee that the converged pattern will minimize overall dissonance (a "global optimum"). We generally refer to the ability of these networks to solve stress minimization problems through adjusting path weights as "structural learning."

4. From Neural to Social Networks

In a recent article, Nowak and Vallacher (1997) note the potential of these computational networks for modeling group dynamics. This approach promises to provide a fertile expansion to social network analysis, which has often assumed that social ties are binary and static. A neural network provides a way to dynamically model a social network in which learning occurs at both the individual and structural levels, as relations evolve in response to the behaviors they constrain.

In addition to variation in path strength, neural networks typically have paths that "inhibit" as well as "excite." That is, nodes may be connected with negative as well as positive weights. In a social network application, agents connected by negative ties might correspond to "negative referents" (Schwartz and Ames, 1977), who provoke differentiation rather than imitation. However, we postpone consideration of negative ties until after we have explored the effects of allowing the strength of ties to vary. Weights in our model are presently constrained to positive values less than unity. Two nodes with a relatively strong tie represent "friends," or agents who experience a high frequency or intensity of contact. Conversely, two nodes with weights near zero represent "strangers," or agents who will have little contact. We allow for inhibitory effects by setting the binary states of nodes as either $+1$ or -1. Thus, a node with a negative state and a positive weight with another node may inhibit the target, causing it to also adopt a negative state.

5. Model Design

In this study, we use an attractor net to simulate a team of up to 15 stylized agents in a fully connected network of N nodes and $N(N-1)$ weighted ties between nodes.[4] Our model consists of four basic components:

(1) An agent-based stochastic decision algorithm for choosing to work or shirk and to impose positive or negative sanctions on targeted individuals.
(2) A benefit function by which individual decisions translate into outcomes that are evaluated as satisfactory or unsatisfactory relative to expectations.
(3) A reinforcement learning algorithm that modifies choice propensities in response to satisfaction with associated outcomes.
(4) A structural learning algorithm that modifies the strength of ties in response to structural imbalances in the behaviors of connected agents.

5.1. Decision Algorithm

Each of the N agents must make $N + 1$ independent decisions during each round. For the single "compliance" decision, each agent must choose whether to contribute to the public good. For the N "approval" decisions, each must choose independently whether to approve or disapprove of every individual agent in the network. To simplify, we assume agents must choose between only two options for each decision, to work or shirk, and to approve or disapprove. Neutrality (or half-hearted effort) is not a behavioral option but an emergent property when behavior is random.

Note that the N approval decisions include an agent's decision to approve or disapprove of itself. We implement self-approval (which might also be interpreted as "self-esteem") so that agents may respond to disapproval by backing away from the sanctioner rather than altering their behavior. For example, if i has low self-esteem and j disapproves of i, then i and j are in agreement regarding approval of i. However, if i has high self-esteem and j disapproves of i, then they disagree and an unbalanced situation arises. Depending on the array of weights and states at that time, the agents will attempt to resolve the tension by altering their relation (moving farther apart) or by changing their behavior (i might have lowered self-esteem, modify the behavior that earned j's disapproval, or j might stop disapproving of i).

These compliance and approval decisions are stochastic. The decision algorithm assumes each agent i has some propensity P_{is} that determines the probability that i will activate binary state s, where $s = 0$ for the compliance decision and $s = \{1, \ldots, N\}$ for the decision to approve of agent s.[5] Each propensity is a function of i's bias ($-1.0 < B_{is} < 1.0$), plus mimetic pressure from other agents. "Bias" refers to i's "internal compass" that is shaped by individual experience through reinforcement. "Mimetic pressure" refers to implicit social pressures to conform with an emergent norm in a reference group.

Mimetic pressure ($-1.0 < M_{is} < 1.0$) is a function of the states of all other agents j, weighted by the magnitude of the dyadic tie between i and j ($0.0 < w_{ij} < 1.0$). For the compliance decision ($s = 0$), i observes the $N - 1$ other members and aggregates their compliance states ($c_j = \pm 1$), weighting each state by i's attraction to that agent (w_{ij}):

$$M_{i0} = \frac{\sum_{j=1}^{N} w_{ij} c_j}{N - 1}, \quad j \neq i \tag{1}$$

To the extent that members of i's reference group (other agents to which i is strongly tied) agree on a given behavior, i perceives an emergent norm to engage in that behavior. The higher the agreement, the stronger the social pressure on i to comply with the emergent norm. This applies also for approval, as i's mimetic pressure to approve of k ($s = k$) is a sum of all j's decisions to approve of k ($a_{jk} = \pm 1$), weighted by i's attraction to j, where j is any of the $N - 2$ agents other than i and k:

$$M_{ik} = \frac{\sum_{j=1}^{N} a_{jk} w_{ij}}{N - 2}, \quad j \neq i, k \tag{2}$$

(Note that for self-approval, $k = i$ and j can be any of $N - 1$ other agents.)

Even strong social pressure does not guarantee that an agent will imitate its friends, however. “Mimetic pressure” translates into “mimetic influence” only to the extent that the receiving agent is dependent on the group for normative guidance. Let δ_i represent i's normative dependence, ranging from 0.0 (i makes decisions based entirely on its reinforcement history) to 1.0 (i seeks only to imitate friends).

The resulting behavioral propensity is then the sum of i's state bias (B_{is}) and mimetic pressure on i to conform (M_{is}), weighted by δ_i:

$$P_{is} = (1.0 - \delta_i) B_{is} + \delta_i M_{is} \tag{3}$$

where P_{is} represents i's propensity to activate any of its $N + 1$ states and $-1.0 < P_{is} < 1.0$.

5.2. *Benefit Function*

The model assumes that the propensity for any given state (e.g., work or approval) is modified by the associated costs and benefits. In our task-group application, the cost of hard work may be offset by two types of benefits: a higher group wage (or team bonus) and social approval by one's peers. All agents receive an equal share of the group wage, regardless of contribution. Let C represent this individual share, or $1/N$th of the sum of individual efforts ($c_i = \pm 1$):

$$C = \frac{1}{N} \sum_{i=1}^{N} c_i \tag{4}$$

The second source of benefit is social support from one's peers (A_i). Support is highest when i receives approval ($a_{ji} = 1$) from all of its friends ($w_{ij} \approx 1.0$), where $a_{ji} = \pm 1$ and $i \neq j$:

$$A_i = \frac{1}{N} \sum_{j=1}^{N} a_{ji} w_{ij} \tag{5}$$

Both C and A_i thus vary in the range (−1.0, 1.0). Agents in this model use the origin as a point of reference. Negative levels of group wage or social support will reduce their satisfaction and positive levels will increase their satisfaction.

The model assumes that agents weigh the benefits from work against the associated costs. Following Coleman, we assume that approval is costless or nearly so (1990, p. 277), but work involves costly effort (e). This cost of working is constrained by two assumptions, that loafing is more cost-effective than working ($1/N < e$) and that the group realizes a Pareto optimal collective benefit when everyone contributes ($e < 1$). Within these limits, $e = 0.2$ provides a useful calibration, given the range of group sizes that we are investigating and the relative importance we have assigned to the group wage.[6] To study the effect of dependence on compliance, we follow Flache and Macy and use a cost level that makes collective action not so difficult that sanctions cannot help and not so easy that sanctions are never needed.

The evaluation of the previous round's decision to work or shirk takes into account the agent's share of the total group wage, the cost of working, and the value of approval associated with the work decision. Instrumental (γ) and affective (β) dependence together determine the relative importance of wage and approval in evaluating satisfaction with the previous decisions to comply:

$$S_{i0} = \frac{\gamma_i C + \beta_i A_i - ec_i}{\gamma_i + \beta_i + e} \tag{6}$$

where S_{i0} is i's satisfaction with the compliance decision $(s = 0)$, γ_i is i's instrumental dependence, β_i is i's affective dependence, and C, A, e, and c are as defined in Eqs. (4) and (5). The denominator simply norms S_{i0} to less than unity in absolute value.

The decision by i to approve of j is evaluated in the same way except that the collective action problem is now disaggregated into $N(N-1)$ interactions, one for each of the directed ties between agents. Rather than taking into account overall group effort and overall approval received from the group, i considers only i's benefit from j's effort and the approval received from j, or:

$$S_{ij} = \frac{\gamma_i c_j + \beta_i a_{ji} w_{ij}}{\gamma_i + \beta_i} \tag{7}$$

(Note that self-approval has no reinforcement consequences.)

Clearly, satisfaction may take positive or negative values. In this model, agents use the origin as a point of reference. They find levels of S_{i0} or S_{ij} above 0.0 to be satisfying and find levels below 0.0 to be dissatisfying.

5.3. *Reinforcement Learning*

Agents' behavior and relations change over time through two types of learning: reinforcement and structural. Reinforcement learning increases agents' propensities toward states associated with positive evaluations. Structural learning strengthens ties (or reduces social distance) between nodes that simultaneously adopt similar states.

In an attractor network, reinforcement learning can be used to modify thresholds at which agents change states in response to other connected nodes. Reinforcement allows agents to seek configurations of states that locally maximize their own satisfaction, regardless of the effects this might have on structural imbalance in relations.

Reinforcement learning is based on Thorndike's (1932) "Law of Effect." At time $t+1$, each agent retrospectively examines the $N+1$ choices it made at time t, evaluates the outcomes that were associated with these choices (formalized in Eqs. (6) and (7)), and then modifies the bias toward repeating those acts that increase subjective rewards and decrease subjective punishments. Agent i's bias toward compliance at time t ($B_{i0,t}$) is reinforced when current work effort seems to pay ($c_{it} = 1$ and $S_{i0,t} > 0$) or when shirking is costly ($c_{it} = -1$ and $S_{i0,t} < 0$), causing the bias to increase in either case:

$$B_{i0,(t+1)} = B_{i0,t} + S_{i0,t}c_{it}(1 - B_{i0,t}) \tag{8a}$$

Conversely, if feckless behavior pays off ($c_{it} = -1$ and $S_{i0,t} > 0$) or hard work is suckered ($c_{it} = 1$ and $S_{i0,t} < 0$), then the bias toward work is decreased:

$$B_{i0,(t+1)} = B_{i0,t} + S_{i0,t}c_{it}B_{i0,t} \tag{8b}$$

The bias toward approval is modified in the same way. Agent i's bias toward approval of j ($B_{ij,t}$) increases when approval has a satisfactory outcome ($a_{ij,t} = 1$ and $S_{ij,t} > 0$) or when disapproval is costly ($a_{ij,t} = -1$ and $S_{ij,t} < 0$):

$$B_{ij,(t+1)} = B_{ij,t} + S_{ij,t}a_{ij,t}(1 - B_{ij,t}) \tag{9a}$$

Otherwise, the bias decreases:

$$B_{ij,(t+1)} = B_{ij,t} + S_{ij,t}a_{ij,t}B_{ij,t} \tag{9b}$$

As we assume that agents always have a $+1.0$ weight toward themselves, we also assume that agents always have a $+1.0$ bias to approve of themselves. Thus they will only have low self-esteem if they are "convinced" by overwhelming disapproval from friends.

5.4. *Structural Learning*

Structural learning is the tendency of the network to seek configurations that represent a local energy minimum. In this application, we use a simple Hebbian learning rule corresponding at the agent level to the principle of homophily: agents grow closer to other agents that are similar to themselves. Weights between agents start with random values from 0.0 to 1.0 (exclusive) and then adjust synchronously, according to perceived similarity.[7]

We assign agents three sets of decisions by which they judge their similarity with other agents: *compliance*, *self-approval*, and *other-approval*. In computing their perceived level of similarity, agents weight these three decision sets by their own propensity to perform each action. We thus take these behavioral propensities to represent an agent's underlying "confidence" in that decision, where highly confident decisions ($P_{is} \approx \pm 1.0$) are weighted more heavily in judging similarity than are ambivalent or random decisions ($P_{is} \approx 0$). We also assume that agents are aware of their own behavioral propensities but do not have access to this dispositional information about others. This reproduces an inherent information asymmetry found in the social world. We know our own intentions and inclinations, even if our observed behavior is a "fluke" (an accident or experiment), but we have no direct access to this information about others. We can observe others' behavior but not their intentions.

In our implementation, agent i compares its own vector of behavioral propensities against the observed behaviors of all other agents j. The assessment yields a score for i's perceived similarity with j ranging from $+1.0$ (i is 100% confident in all decisions and j makes all the same choices) to -1.0 (i is 100% confident in all decisions and j makes all the opposite choices). Based on its relative similarity or dissimilarity with j, i will adjust w_{ij}, its "attraction" to j:

$$\Delta w_{ij} = \frac{P_{i0}c_j}{3} + \frac{1}{3(N-2)} \sum_{j=1}^{N} P_{ik}a_{jk} + \frac{P_{ii}a_{ji}}{3}, \quad j \neq i, k \tag{10}$$

where w_{ij} is i's attraction to j, P_{is} is i's propensity for compliance ($s = 0$) and approval ($s = k$), c_j is j's decision to comply, and a_{jk} is j's approval of k (not including j). For increasing weights,[8] the weight at time $t + 1$ is then:

$$w_{ij,t+1} = w_{ijt} + \Delta w_{ijt}(1 - w_{ijt}) \tag{11}$$

The weight between two agents with maximally different states will approach 0.0 asymptotically over time and the weight between two agents with persistently identical states will approach 1.0.

6. Results

In this study, we measure the levels of cooperation in groups that vary in size and in extent of mimetic influence. In a series of simulations, we manipulate the size of the group (from 3 to 15 in unit increments) and the mean of normative dependence (from 0.0 to 1.0 in increments of 0.05). Individual agents' δ_i scores are randomly distributed about the target mean.[9] We do not manipulate instrumental and affective dependence in this study, but will do so in future research, comparing simulated groups with varying levels of γ and β. Here, we assume that γ and β are uncorrelated and normally distributed around an expected value (0.5) that is constant across all experimental conditions. Therefore, the expected values for γ and β are equal, which means that, on average, the group wage and approval weigh equally in determining aggregate satisfaction with the compliance and approval decisions.

Our simulation experiments map the coordination of collective action across a parameter space of normative dependence and group size. This 13 by 21 (group size by normative dependence) factorial design yields 273 experimental conditions. We replicate each of these conditions 50 times for a total of 13650 simulations. Each trial runs for 5000 iterations or until convergence, which we define as the absence of change in the compliance decision by any agent for 100 consecutive iterations.[10]

The simulations reveal an interesting curvilinear effect of mimetic influence, evident in figure 1 as normative dependence (δ) increases from 0.0 to 1.00, at all levels of group size (from $N = 3$ to 15).

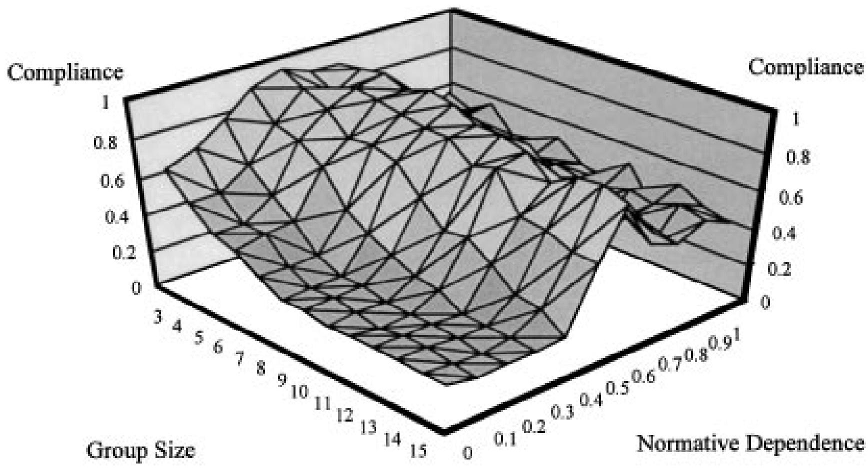

Figure 1. Compliance by group size (N) and normative dependence (δ).[12]

For all levels of N, compliance is relatively low at $\delta = 0$, where agents are acting entirely out of adaptive self-interest. In order for a group to have high compliance at this level of δ, they must coordinate a critical mass of simultaneous cooperators, or a number sufficient to make compliance with work norms rewarding to workers. Imitation simplifies this coordination of concerted action, as indicated by the increasing levels of compliance as δ increases. However, very high levels of δ have the opposite effect on compliance. If agents place a much higher priority on imitating their friends than on repeating successful behaviors, they lose the ability to learn collectively beneficial solutions. Further, as normative dependence reaches high levels, agents' imitation of their friends' approval choices undermines their use of approval for social control. At the maximum of δ, the agents become "lemmings" who adjust propensities only through mimetic influence, and thus their behavior is effectively random regarding the collective action problem.

This curvilinear effect of δ is evident at all group sizes we tested, from $N = 3$ to 15. However, the effect of group size depends on the level of normative dependence. At low levels of δ, compliance declines with increasing group size, an effect that is consistent with most collective action theory (Olson, 1965). This reflects the increasing coordination complexity of reaching a high-compliance solution as group size increases. Adaptive agents are much more likely to reach high levels of mutual compliance (and to then find it rewarding) in small groups.

In the middle range of δ, where compliance is highest, group size has the same negative effect on compliance, but it also influences the width of the compliant plateau. Though each group size has some threshold of δ that will facilitate a high level of compliance, this threshold rises as group size increases. For example, due to the relative simplicity of coordinating compliance in small groups (e.g., $N = 3$), normative dependence of about $\delta = 0.3$ suffices to allow the group to attain a high level of compliance. However, in relatively large groups (e.g., $N = 15$), nearly twice as much normative dependence is needed to

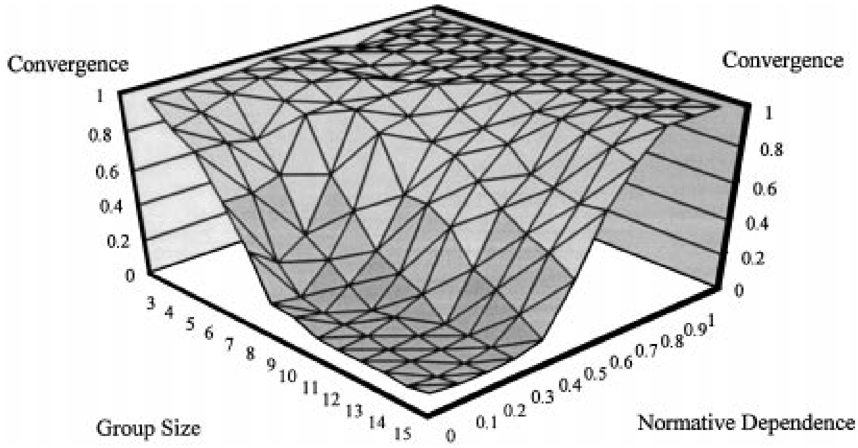

Figure 2. Convergence by group size (N) and normative dependence (δ).

promote widespread compliance. This results in a much narrower ridge for larger groups, as compliance falls off abruptly at higher levels of δ. Of course, group size is irrelevant to compliance in the very high range of δ, as lemmings are not responsive to either the cost or benefit of working.

The pattern of convergence across the parameter space also shows how normative dependence and group size influence levels of compliance by altering coordination complexity. As figure 2 shows, the network converges on a stable solution easily when normative dependence is high but the convergence rate drops steeply as reinforcement learning dominates mimetic influence. Imitation thus allows the network to simplify the coordination problem across all group sizes, but this simplification only improves compliance when conformity is tempered by the lessons of experience.

The size of the group also has the predicted monotonic effect on convergence. As noted above, coordination of satisfactory high-compliance solutions becomes difficult as group size increases. Because agents are dissatisfied with low-compliance solutions, the network remains in flux, unable to find a stable solution within 5000 iterations. Within this range of group size, N seems somewhat less important than δ, as even the smallest groups do not always converge without any imitation, and high imitation allows even groups of $N = 15$ to converge.

Group cohesion (the mean of all weighted approval values) shows a somewhat different pattern, but also supports earlier work on the "weakness of strong ties." At maximum δ, approval behavior is essentially random, so cohesion remains at the midpoint (approximately 0 on a scale of -1.0 to $+1.0$). Across the middle region of δ, where compliance reaches its peak, cohesion is also quite high, as agents are rewarding each other's hard work with approval. However, at very low δ, where agents are adaptive learners rather than imitators, approval remains quite high even though compliance has plummeted. As Flache and Macy (1996) have shown, the groupwise exchange of compliance for approval has much

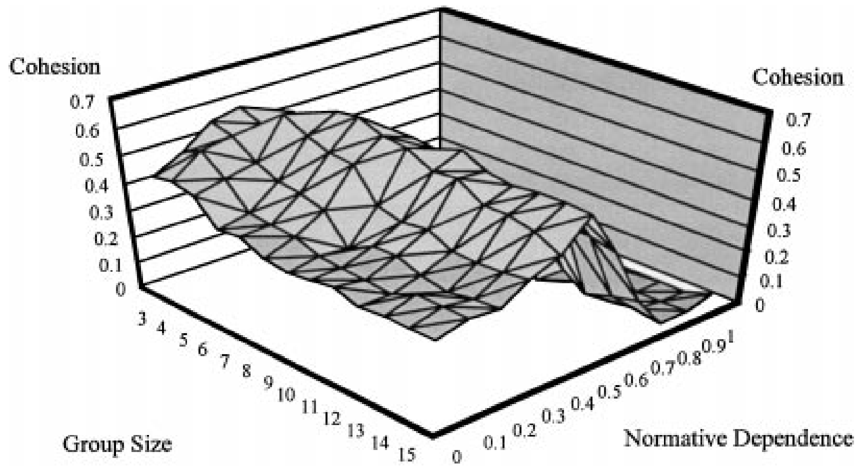

Figure 3. Cohesion by group size (N) and normative dependence (δ).

higher coordination complexity than does the pairwise exchange of approval for approval. Among self-directed (low δ) adaptive agents, social pressures therefore tend to flow into the maintenance of personal relationships rather than enforcement of compliance with group obligations.

7. Conclusion

This study demonstrates a theoretical tool that may shed new light on central questions of group processes and informal social control. Previous work modeling social dilemmas among networks of autonomous adaptive agents has not allowed structural learning (changes in the strength of social ties) or mimetic influence between friends. In effect, these previous models have been limited to the special case of $w_{ij} = 1.0$ and $\delta = 0.0$.

In contrast, work in attractor neural networks has modeled changes in node states as a function of the states of other nodes and the strength of ties connecting nodes. However, these studies have not allowed agents to learn from individual experience. In effect, these models have been limited to the special case of $\delta = 1.0$.

We combine structural and reinforcement learning in order to explore the full range of mimetic influence in a dynamic social network. In our neural network application, learning occurs at both the individual and structural levels, and influence occurs through both sanctions and imitation. Nevertheless, the model maintains a reasonable parsimony at the level of individual agents. The evolution of relations and decisions in the base model derives from two simple and broadly supported theories, the Law of Attraction (for structural learning) and the Law of Effect (for reinforcement learning). To this we add an equally simple tendency to imitate similar others.

Structural learning provides a particularly promising avenue for future research. Allowing the strength of ties to vary as a function of agents' behavioral similarity poses a possible pitfall for coordination of informal social control. With variable weights, agents may distance themselves from others who sanction them, rather than complying with a productivity norm. In this way, cohesive cliques may form, which exchange approval internally and imitate one another in shirking, while also disregarding punishment by other members of the work group. The formation of such a “deviant subculture” poses an additional coordination problem for a system-level cooperative solution. When agents are unable to influence members of oppositional cliques, the network should have a harder time reaching critical mass, even at optimal levels of mimetic influence.

In fact, we rarely observed such cohesive subgroups in these experiments, and never observed them in groups with high levels of normative dependence. We hypothesize that the formation of such oppositional cliques in a population of imitative agents requires the potential for negative ties.[11] In our model, weights can only vary in the interval 0.0 to 1.0. Without negative ties, agents cannot differentiate. Suppose two cliques formed that were maximally distant from one another, that is, with tie weights approaching 0.0. Nevertheless, over time, even slight similarities will grow into strong friendships in the presence of strong mimetic influence, due to the positive feedback between conformity and homophily in our model. Growing friendships further bolster behavioral similarities, which in turn strengthen friendships.

In future work, we plan to explore the effects of negative influence on the coordination of collective action. Not only will agents be attracted to similar others, but they will be repulsed by those who are different. Not only will they imitate “friends” and ignore “strangers,” but they will differentiate themselves from “enemies.” With the addition of negative ties (and negative influence), we expect that internally cohesive cliques will develop in groups with relatively high normative dependence. If these cliques are divided on the question of contribution to the public good, this will pose a problem for coordination of informal social control that is not evident in these simulations.

To summarize the key findings, this study shows that moderate levels of mimetic influence may aid collective action by helping the group coordinate a critical mass of cooperators by random walk. However, very high levels of mimetic influence produce “lemmings” who seek agreement even when this results in a failure to achieve group goals. These highly imitative friendship networks thus tend to evolve into a “melting pot,” with agents moving inexorably toward an arbitrary consensus.

We believe this dynamic reflects our assumption that social ties can only be positive. In contrast, we expect negative social ties to foster internal differentiation and the creation of stable and internally cohesive subgroups. This hypothesis suggests a curious paradox: Contrary to the pluralist assumption that tolerance of differences is a precondition for the maintenance of diversity, intolerance may in fact promote the preservation of diversity. We leave the investigation of this puzzle to later work.

Acknowledgment

This research was supported by grants from the National Science Foundation, #SBR 98-01691 to the first two authors and #SBR 95-11461 to the second author, and from the German National Science Foundation, No. He 1412/4-1 to the third author. The authors acknowledge Andrzej Nowak, Frans Stokman, Michael Spivey, and Ali Hadi for their constructive comments.

Notes

1. Flache (1996) uses both simulation and game theoretical analysis to derive the "weakness of strong ties." In addition, he presents social-psychological experiments supporting the hypothesis.
2. An executable copy of this model is available for the Windows 95 operating system at: *http://www.people.cornell.edu/pages/jak20/*.
3. For example, Newcomb (1961) found that "high attraction preferences tended to change in favor of individuals with whom they were more closely in agreement." See also Cohen, 1977; Kandel, 1978; and Verbrugge, 1977. We are intrigued by the possibility that neural and social networks may have common homophilous properties and the implications this holds for the existence of "social memory" stored in network configurations. However, we leave the exploration of memory in social networks for future research.
4. Each node also has a reflexive tie. By definition, an agent is always 100% similar to itself, and we therefore fix reflexive weights at $+1.0$.
5. More precisely, probabilities are an approximate sigmoid function of propensities. This makes the model deterministic outside the inflection points of the curve.
6. For computational convenience, we use positive and negative values to represent the binary states of incurring or escaping the cost of compliance. This parallels our use of positive and negative values to represent the binary decision to comply. An agent who complies contributes a positive amount ($c_i = +1$) to the public good but incurs a negative individual payoff ($-e$). Conversely, shirkers subtract from the public good ($c_i = -1$) but enjoy a positive individual payoff (e).
7. Note that the random start means that the weight matrix is initially asymmetric. Further, any two agents may have different perceptions of their similarity, and thus may not update their mutual weights (w_{ij} and w_{ji}) congruently. This differs from Hopfield's (1982) classical model, in which weights are symmetric. In addition, Hopfield notes that there appears to be nothing analogous to simultaneous updates occurring in living neurological systems, and so uses asynchronous random updates. However, our objective is not to emulate neurological systems, and we therefore prefer to begin with the computationally simpler simultaneous-update model and save asynchronous updating for future research.
8. For decreasing weights, the change is multiplied by w_{ijt} rather than $(1 - w_{ijt})$.
9. Note that the distribution of δ_i becomes skewed as we manipulate the group mean away from the midpoint of the parameter space (0.0, 1.0). At the limits ($\delta = 0.0$ and $\delta = 1.0$), the population becomes homogenous.
10. At this equilibrium, given an array of instrumental, affective, and normative dependence scores and a random start, the network has learned a stable solution to the social dilemma. Of course, this "attractor" may not represent a collectively optimal solution in terms of overall group compliance. Also, path weights and approval decisions do not necessarily converge at this point, but we use compliance as the criterion because this is our principal dependent variable.
11. Using a cellular model in which agents may interact only with their four nearest neighbors, Axelrod (1997) shows that subgroup differentiation can occur without negative ties. However, this is not possible in a fully connected graph such as the production teams that we are modeling.
12. This graph represents 13650 runs of the simulation, with each cell representing the 100 replications within a given region of the parameter space. As a further test of robustness, we allowed the distributions of individuals' affective and normative dependence to vary across replications.

References

Axelrod, R. (1997), "The Dissemination of Culture: A Model with Local Convergence and Global Polarization," *Journal of Conflict Resolution*, 41, 203–226.

Blau, P.M. (1977), *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Blau, P.M. and J. Schwartz (1984), *Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations*. New York, NY: Academic Press.

Byrne, D.E. (1971), *The Attraction Paradigm*. New York, NY: Academic Press.

Byrne, D.E. and W. Griffit (1966), "A Developmental Investigation of the Law of Attraction," *Journal of Personality and Social Psychology*, 4, 699–702.

Carley, K. (1991), "A Theory of Group Stability," *American Sociological Review*, 56, 331–354.

Cartwright, D. and F. Harary (1966), "Structural Balance: A Generalization of Heider's Theory," *Psychological Review*, 63, 277–293.

Churchland, P.S. and T.J. Sejnowski (1994), *The Computational Brain*. Cambridge, MA: MIT Press.

Cohen, J.M. (1977), "Sources of Peer Group Homogeneity," *Sociology of Education*, 50, 227–241.

Coleman, J.S. (1990), *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Dawes, R.M. (1980), "Social Dilemmas," *Annual Review of Sociology*, 31, 169–193.

Festinger, L., S. Schachter and K. Back (1950), *Social Pressures in Informal Groups*. New York, NY: Harper.

Flache, A. (1996), *The Double Edge of Networks*. Amsterdam: Thesis Publishers.

Flache, A. and M.W. Macy (1996), "The Weakness of Strong Ties: Collective Action Failure in a Highly Cohesive Group," *Journal of Mathematical Sociology*, 21.

Hebb, D.O. (1949), *The Organisation of Behavior: A Neuropsychological Approach*. New York, NY: John Wiley and Sons.

Heckathorn, D.D. (1988), "Collective Sanctions and the Creation of Prisoner's Dilemma Norms," *American Journal of Sociology*, 94, 535–562.

Heider, F. (1958), *The Psychology of Interpersonal Relations*. New York, NY: John Wiley and Sons.

Homans, G.C. (1951), *The Human Group*. New York, NY: Harcourt Brace.

Homans, G.C. (1974), *Social Behavior: Its Elementary Forms*. New York, NY: Harcourt Brace Jovanovich.

Hopfield, J.J. (1982), "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proceedings of the American Academy of Sciences*, Vol. 79, pp. 2554–2558.

Hopfield, J.J. (1984), "Neurons with Graded Response have Collective Computational Properties like those of Two-State Neurons," *Proceedings of the American Academy of Sciences*, Vol. 81, pp. 3088–3092.

Hopfield, J.J. and D.W. Tank (1985), " 'Neural' Computation of Decisions in Optimization Problems," *Biological Cybernetics*, 52, 141–152.

Kandel, D.B. (1978), "Homophily, Selection, and Socialization in Adolescent Friendships," *American Journal of Sociology*, 84, 427–436.

Lott, A.J. and B.E. Lott (1965), "Group Cohesiveness as Interpersonal Attraction: A Review of Relationships with Antecedent and Consequent Variables," *Psychological Bulletin*, 64, 259–309.

Macleod, J. (1995), *Ain't No Makin It: Aspirations and Attainment in a Low Income Neighborhood*. Boulder, CO: Westview Press.

Macy, M.W. (1990), "Learning Theory and the Logic of Critical Mass," *American Sociological Review*, 55, 809–826.

Macy, M.W. (1991), "Chains of Cooperation: Threshold Effects in Collective Action," *American Sociological Review*, 56, 730–747.

Macy, M.W., J.A. Kitts and A. Flache (1997), "The Weakness of Strong Ties II: Collective Action Failure in a Self-Organizing Social Network," Presented at the 1997 Annual Meeting of the American Sociological Association.

McPherson, J.M. and L. Smith-Lovin (1987), "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups," *American Sociological Review*, 52, 370–379.

Newcomb, T.M. (1961), *The Acquaintance Process*. New York, NY: Holt, Rhinehart and Winston.

Nowak, A. and R. Vallacher (1997), "Computational Social Psychology: Cellular Automata and Neural Network Models of Interpersonal Dynamics," in S. Read and L. Miller (Eds.), *Connectionist Models of Social Reasoning and Social Behavior*, Mahwah, NJ: Lawrence Erlbaum.

Oliver, P. (1980), "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations," *American Journal of Sociology*, 85, 1356–1375.

Olson, M. (1965), *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Quinlan, P.T. (1991), *Connectionism and Psychology: A Psychological Perspective on New Connectionist Research*. Chicago, IL: University of Chicago Press.

Rummelhart, D. and J. McClelland (1988), *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. Cambridge, MA: MIT Press.

Schwartz, S.H. and R.E. Ames (1977), “Positive and Negative Referent Others as Sources of Influence: A Case of Helping,” *Sociometry*, 40, 12–21.

Seashore, S.E. (1954), “Group Cohesiveness in the Industrial Work Group,” Ann Arbor, MI: University of Michigan Institute for Social Research.

Thorndike, R. (1932), *The Fundamentals of Learning*. Columbia University, New York, NY: Teachers College.

Turner, R.H. and L.M. Killian (1957), *Collective Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Verbrugge, L.M. (1977), “The Structure of Adult Friendship Choices,” *Social Forces*, 56, 576–597.

Willis, P.E. (1978), *Profane Cultures*. London: Routledge and Kegan Paul.

Yinger, M. (1982), *Countercultures*. New York, NY: Free Press.

Zeggelink, E. (1994), “Dynamics of Structure: An Individual Oriented Approach,” *Social Networks*, 16, 295–333.

James Kitts is a doctoral candidate at Cornell University, having earned a M.S. in Natural Resources from the University of Michigan and a M.A. in Sociology from Cornell. His dissertation analyzes the mortality of American communes as a function of organizational structure and population dynamics. He is interested in structural models of collective action, social influence, and organizational dynamics.

Michael Macy (Ph.D. Harvard University, 1985) is Professor of Sociology at Cornell University. He studies collective action and social influence in networks, using computer simulation of adaptive agents and laboratory experiments with human subjects. A new project with David Strang models diffusion of management fads across firms. For more information, please visit his website at: *www.people.cornell.edu/pages/mwm14*.

Andreas Flache took his master’s degree in Computer Science and Social Sciences from the University of Koblenz, Germany. He conducted his Ph.D. research on the ‘double edge of networks’ at the ICS (Interuniversity Center for Social Science Theory and Methodology) in Groningen, The Netherlands. He took his Ph.D. in 1996. At the moment, he is postdoc at the Department of Philosophy of the University of Bayreuth, Germany. His main research interests concern social dilemmas, formal modeling of collective behavior, and social networks.