Words: 9,535
Date: 7-Apr-98

AGENCY AND INTERACTION*

Peter J. Burke

Washington State University

* An earlier version of this paper was presented at the annual meetings of the American Sociological Association, August, 1996. I would like to thank Jan Stets and Louis Gray for comments on an earlier draft.
Address all correspondence to Peter J. Burke, Department of Sociology, Washington State University, Pullman, WA 99164-4020. E-mail: burkep@wsu.edu. URL: http:// burkep.libarts.wsu.edu

AGENCY AND INTERACTION

Abstract

This paper considers the connection between macro-level group characteristics (social structure) and the micro-level interaction between individuals by focusing on the nature of the individual that is affected by structural characteristics and, with others, produces the patterns of interaction that characterize aspects of social structure. Two basic models of the individual are considered: one based on the backward-looking, passive models of operant learning theory implicit in social exchange theory, the other based on the forward-looking, agency model implicit in identity theory and affect control theory (Smith-Lovin and Heise 1988). Computer simulations are used to "solve" the interactive consequences of these models under a variety of structural conditions, using different formulations of the operant learning model and different agent strategies. The results show that (1) stable interaction structures do not emerge under any forms of the operant learning model, but do under the agent models and (2) the structural conditions are important for the type of structure that emerges for the agent models, but have no impact on the outcomes under the operant models. The implications of these results for sociology and social psychology are discussed.

AGENCY AND INTERACTION

Introduction

In considering the connection between the macro-world of social structures and the micro-world of interaction between individuals, there are two separate issues (Coleman 1990; Turner 1988). The first concerns the nature of the impact of social structure on interacting individuals, while the second concerns the nature of the impact of interacting individuals on the development, change and stability of social structure.[1] Mediating these two processes, however, as shown in Figure 1, adapted from Coleman (1990), lies the interaction patterns evolving between and among individuals who a) have particular natures and b) are embedded in particular social contexts. To the extent that these interaction processes "lock in" to stable patterns, social structure(s) can be said to exist.

(Figure 1 About Here)

As part of an attempt to understand the origins and continuance of social structure, the present paper examines both interaction and the nature of individuals as understood by two fundamentally different orientations that may be characterized as "forward looking" and "backward looking." Identity theory, representing the forward looking model, sees individuals as agents whose behavior is shaped by the individual's trying to achieve or accomplish internally held goals within the constraints of the social system in which they find themselves (Burke 1991, 1997). The second view, represented by behavioral sociology (Burgess and Bushell 1969), exchange theory (Homans 1961; Emerson 1969), and

[1] It is, of course, recognized that "social structures" exist at multiple levels of abstraction and inclusion.

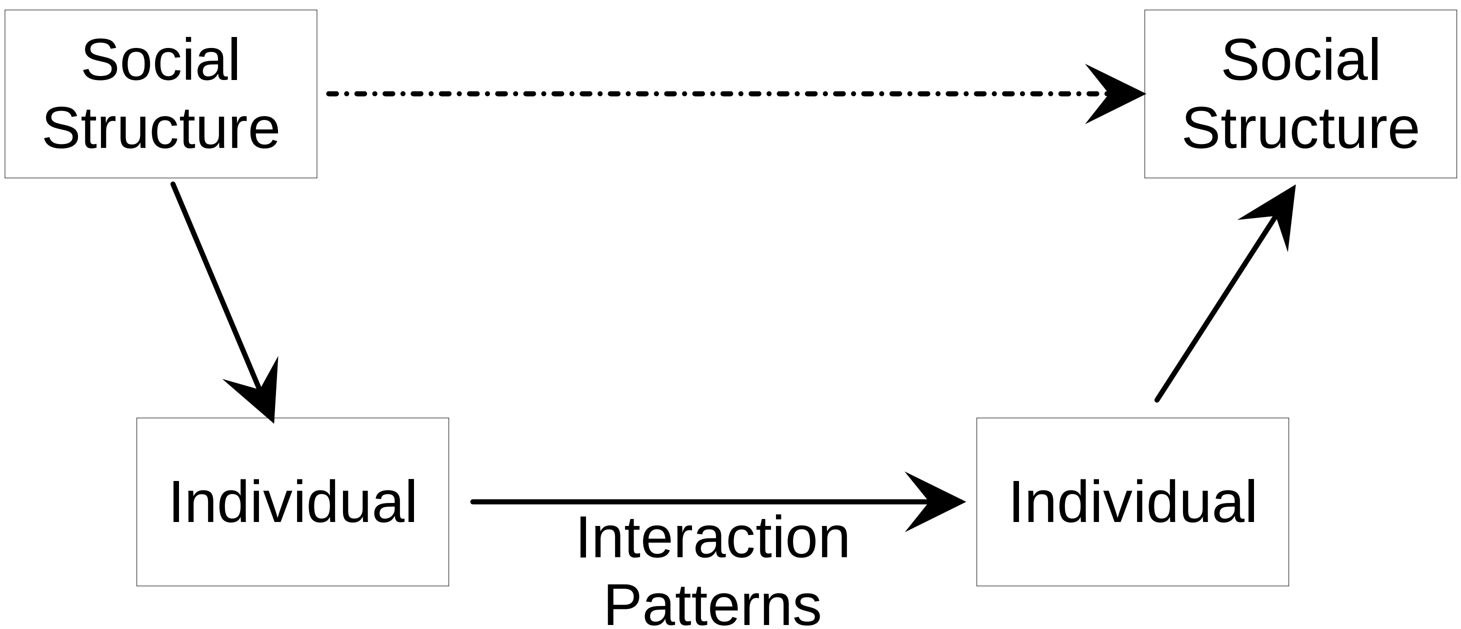

Figure 1. Social Structure and the Individual, adapted from Coleman (1990)

operant learning theory (Estes 1957; Herrnstein 1970) in their pure forms, sees individuals as entities whose behavior is shaped by their history of reinforcement.

The primary difference between these theories lies in the relationship between reinforcements and behavior. In reinforcement theory, behavior occurs "because of …," in identity theory, behavior occurs "in order to…". Since the behavior of individuals, and the interaction process evolving from that behavior, are continuous, as are perceptions of outcomes such as reinforcements, it is often difficult to see whether behavior is the cause of the reinforcement (forward looking) or the reinforcement is the cause of the behavior (backward looking). In addition, there is the question of what difference it makes. Clearly, if it makes no differences for any outcomes of interest, then the question is moot.

In order to begin to answer these questions, I create several simulations that model the process of interaction in each of the two ways (forward looking and backward looking). I then examine the outcomes of these simulations and compare them with each other in order to understand the theoretical implications of the models. At this point, the goal is not to create a model that mimics human behavior, but to explore different theoretical assumptions made in sociology and social psychology, so that agreement can eventually be reached about assumptions that are appropriate to modeling human behavior. Each of the models is explored under a variety of social structural conditions. I do this in order to draw out the implications for our understanding of the impacts of social structure both on the behavior of individuals and the interaction patterns that develop among individuals within particular social structural arrangements.

Theoretical Models

Reinforcement Theory: Behavior Because of…

According to the reinforcement model of operant learning theory, people act on the basis of their history of rewards (reinforcements) and costs (punishments) incurred as a result of the behaviors or operants they have produced in the past (Skinner 1974). This view of behavior lies at the base of exchange theory as proposed by Homans (1961) and Emerson (1969). It is at the core of power-dependence theory (Emerson 1962; Cook and Emerson 1978; Cook and Yamagishi 1992), the satisfaction balance model (Gray and Tallman 1984, 1986, 1987), and a number of other approaches in social psychology (e.g., Thibaut and Kelley 1959). The basic idea is contained in the "law of effect" (Herrnstein 1970) which tells us that the probability of a behavior that is followed by a rewards is increased, while the probability of a behavior that is followed by a punishment or cost is lowered.[2]

The "matching law" (Estes 1957) provides us with a formula for the instantaneous change in the probability of a particular behavior when it is reinforced (or not). Under the matching law, when there are two exclusive alternative behaviors, call them A and B, the probability of behavior A will change toward matching the probability of the reinforcement of A according to the formula

[2] Actually, in Skinner's original formulation, punishment was not considered. Rather, negative reinforcement "strengthens any behavior that reduces or terminates [the stimulus]" (1974:46). Thus, punishment may be seen not as reducing the probability of the punished behavior, but as increasing an alternative behavior that removes the punishment (a behavior which may not have been anticipated by the punisher).

$$P_{n+1} = (1-\theta)\cdot P_n + \theta \qquad \text{if A is reinforced, and} \qquad \text{(Eq. 1)}$$

$$P_{n+1} = (1-\theta)\cdot P_n \qquad \text{if A is not reinforced,} \qquad \text{(Eq. 2)}$$

where P_n is the probability of the behavior A at time n, and θ is the instantaneous magnitude of impact (a constant, between 0 and 1, for a particular learning situation and usually small, on the order of 0.1) that a reinforcement has on the probability of the behavior. From this it can be seen that if choice A is reinforced with probability $p(R_A)$, in the long run choice A will be made with a probability, p(A), that is equal to $p(R_A)$. Thus, $p(A)=p(R_A)$.

This formulation does not take into account the magnitude of reinforcement or punishment, as it only treats choices as either right or wrong. The satisfaction-balance model, proposed by Gray and Tallman (1984) does take the magnitude of reward (and punishment) into account. This model (in its simplest form) suggests that the probability of choice A (in two choice situations) is a function of both the magnitudes and probabilities of rewards as well as the magnitudes and probabilities of punishments contained in the value and cost outcomes according to the formula.

$$p(A) = \frac{V(A)^{1/2} \cdot C(B)^{1/2}}{V(A)^{1/2} \cdot C(B)^{1/2} + V(B)^{1/2} \cdot C(A)^{1/2}}. \qquad \text{(Eq. 3)}$$

V(A), the value of choice A, is a product of the likelihood of A being correct, $p(R_A)$, and the magnitude of the reward for choosing A, (A_r), and of C(B), the cost of choice B, is a product of the likelihood of B being incorrect, 1-$p(R_B)$ and the magnitude of the cost for choosing B, (B_p). The cost of choice A, C(A) and the value of B, V(B), are similarly defined. This formulation only gives the long run expected outcome, rather than the instantaneous change in probability as the results of a prior reinforcement or punishment. The instantaneous formulation (for disjunctive outcomes) may be given as

$$P_{n+1} = (1-\theta)\cdot P_n + \theta \qquad \text{if A is reinforced, and} \qquad \text{(Eq. 4)}$$

$$P_{n+1} = (1-\eta)\cdot P_n \qquad \text{if A is not reinforced,} \qquad \text{(Eq. 5)}$$

where $\theta = k_1\cdot(A_r\cdot B_p)^{1/2}$ is the scaled magnitude of the impact when the choice is rewarded or correct (k is the scaling constant), and $\eta= k_2\cdot(A_p\cdot B_r)^{1/2}$ is the scaled magnitude of the impact when the choice is punished or incorrect. As can be seen, if the magnitudes of the reward and punishment are equal, this simplifies to the matching law. Note that the impact coefficient in the satisfaction-balance model is a variable function of the magnitude of reward and/or punishment rather than a constant as in the matching law.

To summarize, operant learning theory suggests that behavior is probabilistically produced, and that the probability of any particular behavior is a function of the immediately prior probability of that behavior as modified by the most immediate past reward (or punishment). In this way the probability of a behavior contains the *entire relevant reinforcement history of the particular behavior*.

Identity Theory: Behavior in Order to…

According to identity theory, people act as if they are trying to achieve goals. There is, however, no teleology. As illustrated in Figure 2, identities are composed of hierarchical sets of control systems, each of which has four principal parts. First is a standard that holds the self-meanings of the identity along whatever dimensions are socially relevant for defining that identity. Second are perceptions of self-relevant meanings in the current social situation, which perceptions are along the same dimensions of meaning that are coded into the identity standard. Third is a comparator that compares the self-relevant situational perceptions and the identity standard and emits an error signal, which is the difference between the perceptions and the standard. Fourth is the output, which is a function of the error signal emitted by the comparator. For the lower level control system, the output behavior acts upon the social situation to change the self-relevant meanings in the situation, which, in turn, changes the perceptions of those meanings. This process of selecting alternative behaviors continues until there is a match between the perceptions and the standard, at which point the error signal is zero and there is no further behavior change.[3] In this way the input of self-relevant perceptions are controlled. The output of the higher-level control system is the standard for the lower level system.

(Figure 2 About Here)

In identity theory, "goals" are self-meanings that are contained in identity standards, and they are "achieved" when current perceptions of self-relevant meanings in the situation correspond with the self-meanings contained in the identity standards. As pointed out by Freese and Burke (1994) these self-relevant meanings may be symbolic or sign meanings, the latter of which are meanings tied directly to resources in the situation, while the former may be thought of as tied to potential resources that are not present in the situation.

---

[3] In point of fact, there may never be perfect match because exogenous events in the situation act to further disturb the relationship between perceptions and standard.

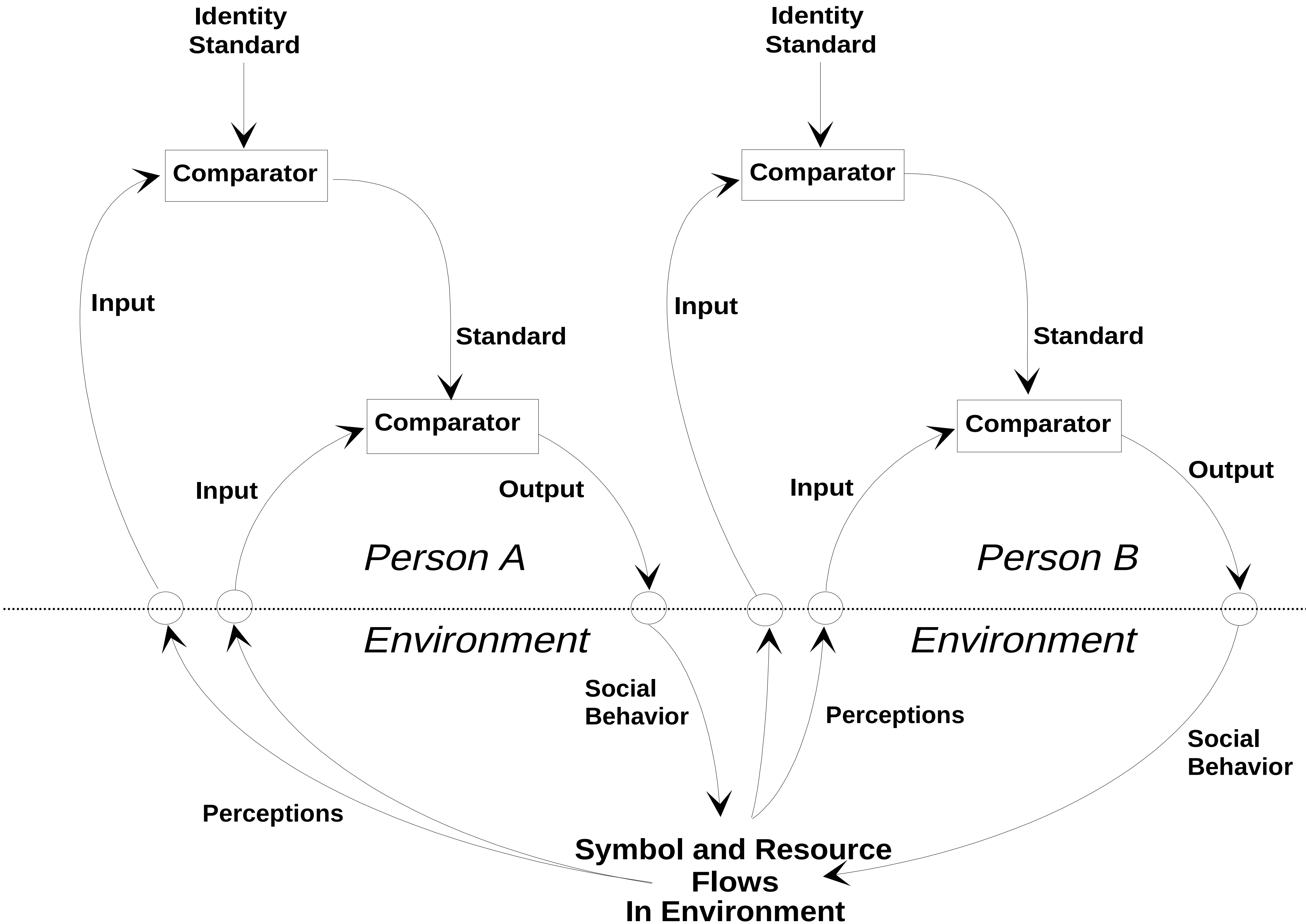

Figure 2. Identity Models for Two Interacting Persons

Actors and Interaction

Any model or theory of an actor must always consider, from the actor's point of view, the question of what to do next, i.e., given the current states of the actor and the situation, what is the next behavior? Operant learning theory and identity theory each provide different answers to this question, and, as stated above, it is the purpose of this paper to explore some of the implications and consequences of these differences. This is not difficult to do for a single individual in a fixed situation. However, if, as in this case, the implications and differences that are of interest are those that result from several actors interacting in a situation, a much more complicated problem exists.

Each of the actor's behaviors has consequences for the behavior of other actors in the situation, which other's behavior, in turn, has consequences for the actor in question. Such self-referencing feedback loops and the contingent nature of the interaction result in outcomes that are highly non-linear and relatively unpredictable with normal analytic techniques. What will be the long run behavior of the individual actors? Will the interaction among the actors become patterned and stable over time? What effects will the underlying models (operant learning theory and identity theory) have on those patterns? These are questions that can only be answered by simulating the actors in an interactive setting (or at least simulating those relevant aspects of the actor that are of theoretical interest).

In doing this I need to further consider the nature of interaction. In the simplest case, interaction occurs when there are two or more individuals in the same situation and the behavior of each influences that of others in the situation. Interaction becomes interesting when the behavior of each of the persons involved becomes patterned and repetitive in relation with the behavior of the others. When this happens, we have the emergence of social structure, or, in the language of complexity theorists, we find attractors for the behaviors of the interactants (Nowak and Lewenstein 1994).

In the present paper, the nature of the influence of one actor on another is the point in question. Emerson (1969), whose work forms much of the basis of exchange theory today, built his theory of interaction and exchange on the operant psychology theory of reinforcement. In his view, an exchange relation exists when, in a two-actor situation, the actions of actor A reinforce the actions of actor B, and the actions of actor B reinforce the actions of actor A. The actions of each party are contingent upon the actions of the other, and each is dependent upon the other for rewards. Emerson did not address the question of how such a relationship might come about, but rather pursued more sociological questions concerning the nature and consequences of the relationship once it has been formed.

For example, he pointed out that power is the obverse of dependence. A has power over B to the extent that B depends upon A for rewards (and vice-versa). Emerson suggests that the main restraint of A's use of power over B is A's own dependence upon B. An imbalance of power promotes power use (thereby altering the exchange ratio within the relation), and balance is a stable state discouraging its further use (Emerson 1969). As a result of this tendency, unbalanced exchange relations tend to change toward balance across continued transactions in the relation. "To have a power advantage is to use it, and to use it is to lose it" (Emerson 1969:391).

Thus, for Emerson, the distribution of reward and punishment potential (power) across a set of interactants sets up patterns of reinforcement and probabilities of behaviors which structure behavior within the group or network in particular ways. These patterns tend, over time, to shift along predictable paths according to the principles of a model that focuses on how those rewards and punishments are used.

Identity theory sees the nature of the influence of one actor on another in a different way. Each actor is an agent whose identity consists of the "goals" they hold for themselves.[4] These goals are the identity standards used to evaluate perceptions of self-relevant meanings arising in the situation. Given goals to seek rewards and avoid punishments, identity theory suggests that people will act in ways that result in a match of perceptions with standards and will alter their behavior only until that match occurs. In the absence of an ability to bring about the match, the standards will change as the result of a higher level control processes, and actors will continue to seek a match between perceptions and (now altered) standards (Burke 1991, 1997; Burke and Cast 1997).

In this view, people act, not on the basis of what has happened to them (perceptions of other's actions as rewarding or punishing) as in the case of the operant model and exchange theory, but on the basis of a comparison of those perceptions with their goal. When these are different, people will change their behavior and an attempt to bring the perceptions and goal (standard) into agreement. Patterns of interaction are achieved and remain stable to the extent that sets of behaviors can be changed to provide congruence between the perceptions and standards of all the interactants.

[4] Again, these goals are nothing more than the sets of meanings (sign and symbolic) that persons holds for themselves as their identity standards. The accomplishment of the goals is nothing more than the perception that these meanings are realized in the current situation (perceptions match standards).

Procedures

Interaction Setting

In order to test and compare these two different theoretical perspectives of the individual, each will be built into a computer model and used in controlled simulations of series of interaction settings. Eight different settings for the simulated interaction are built around a four actor exchange network adapted from the settings used by Molm (1989) and illustrated in Figure 3. In this network, there are four actors, P1 (person 1), P2 (person 2), O1 (other 1), and O2 (other 2). Each actor can act in one of four ways as the result of two binary choices. The first choice is of one or the other exchange partner to whom they are connected (choosing "person" or "other"), the second is to reward or punish that actor. There is no negotiation or bargaining in this setting and the amount of the reward or punishment that can be given is fixed by the experimental design as an unchanging characteristic of the "social structure" of the group.

(Figure 3 About Here)

At each round of exchange, each actor is free to apply rewards and punishments as they will, but each actor is dependent upon others for the rewards and punishments they get, so what one gets may be indirectly a function of the way in which others are treated. The problem for the actors (if they have goals) is to find the treatments of others that yield the largest payoff over the series of exchanges, that is the most rewards and the few-

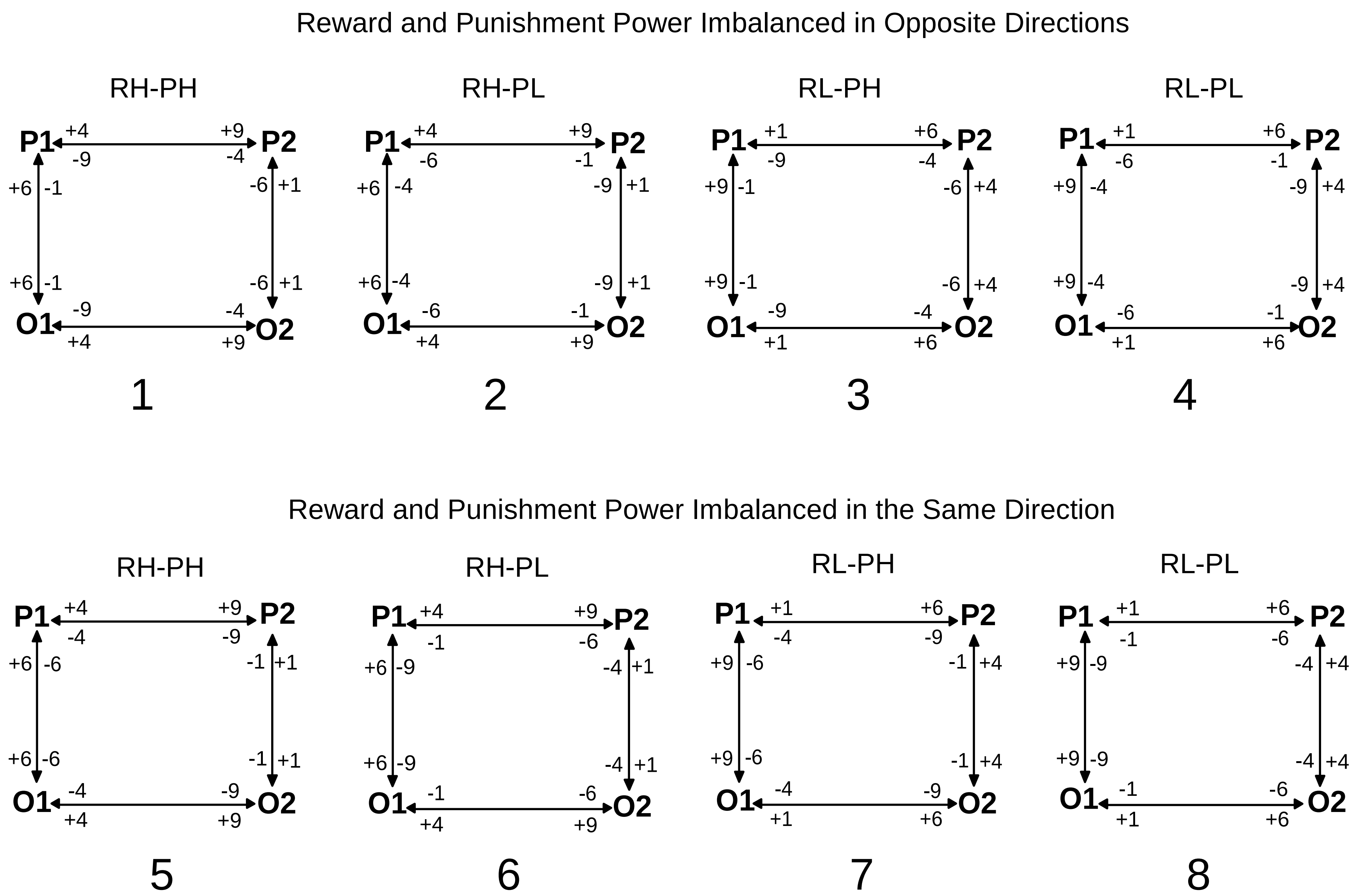

Figure 3. The eight power structure conditions.

est punishments.[5] The total number of exchanges in a session is set at 250. Also, 250 different four-actor simulated groups are run for each of the eight conditions.

Following the original work of Molm (1989), the eight different experimental conditions are set by varying the magnitude of the reward (reward power) and punishment (punishment power) that each can provide others with whom they "exchange." In each of the designs illustrated in Figure 3, the numbers closest to each actor on the arrows are the amounts that they can receive from the actor from whom the arrow is directed. Thus, in Figure 3, condition 1, actor P1 can receive 4 points from P2 or 6 points from O1, and can have 9 points taken away by a punishing move from P2 or 1 point taken away by a punishing move from O1. Each of the other amounts are interpreted in like fashion.

Points that are given in reward or taken away in punishment come from or return to a pool of points held by the experimenter. Thus if O1 takes points away from P1 by a punishing move, O1 does not gain the points P1 only looses them. If O1 gives points to P1, O1 does not loose them P1 only gains them. The costs of providing rewards and punishments to others, therefore, lies in the *indirect* effect those rewards and punishments have in possibly being reciprocated by the others to whom they are directed. It will be noted in Figure 3 that actors P1 and O1 are counterparts of each other as are actors P2 and O2. For this reason I will look only at the behavior of actors P1 and P2.

Each of the eight conditions is labeled by the reward and punishment power of actor P1 relative to actor P2. In half the conditions actor P1 has high reward power and half P1

[5] I am assuming that is the goal. In reality, people may, for example, only want the largest payoff a certain proportion of the time, or may desire a certain level of average payoff.

has low reward power. In half the conditions actor P1 has high punishment power and half P1 has low punishment power. Finally, in half of the conditions the punishment power is in the opposite direction as reward power (actor P1 has high reward power, actor P2 has high punishment power) and half are in the same direction (actor has high reward power and high punishment power). In all conditions, actor P1 has more reward power than actor P2. These conditions are all given in Figure 3.

The Models

The Operant Learning Models. As mentioned above, each actor has a repertoire of four behaviors in which he or she may engage. In each of the five operant learning models that will be considered, each of these possible behaviors has an initial probability of occurrence which is modified by the experience of the actor over time. Rewarded behaviors have their probabilities increased, while punished behaviors have their probabilities decreased, each according to the formulae given in equations 1-5. For the satisfaction-balance model, the magnitude of reward and punishment are given by the number of points received or taken away. If, on a given turn no points are given or taken away, the impact parameters (θ and η in equations 4 and 5 above) take on the value of zero and no change in probabilities occurs.[6] For these models, receiving points in an exchange is a reward, while loosing points is a punishment.

---

[6] Because there are no fixed magnitudes or probabilities for the rewards and punishments associated with particular behavior choices in the present conditions, the impact parameters θ and η are both estimated as $k^*|A_r-A_p|$ and applied to increasing the probability of a "rewarded" behavior ($A_r-A_p > 0$) using Eq. 4 and decrease the probability of a "punished" behavior ($A_r-A_p < 0$) using Eq. 5. As before, A_r and A_p are the

For the basic operant model that does not consider the magnitude of reward or punishment, three submodels are considered which vary in the way that the "no-response" outcome (neither receiving points nor having them taken away by either potential exchange partner) is handled. For the *avoid punishment* model, the no-response and the reward outcome were taken as representing a success in avoiding punishment. In the *seek reward* model, the no-response and the punishment outcome were taken as representing failure to get a reward. In the *neutral middle* model, the no-response did not change the probabilities of any behaviors. One additional model, *seek improvement*, defined reward and punishment a little differently. In this model, the simulated actor viewed an *increase* in points earned or a *decrease* in points lost (compared to the last trial) as a reward. Punishment was viewed as a *decrease* in points earned or an *increase* in points lost. Again, the no-response outcome did not change the probabilities of any potential behavior.

Because the actions of each simulated actor are simultaneous, there is a one-exchange delay in the implementation of the new probability. That is, in exchange one, a particular act is chosen and used. In exchange two, another act is chosen. The consequence of the act in exchange one is seen following the other's exchange two acts. The probabilities of each type of act are changed preceding the third act from the outcomes contained in the second act. The new probabilities are used to select an act for exchange three, while the outcomes of exchange two are seen and used to modify probabilities of acts for exchange four. This continues, then, for the remaining exchanges.

---

magnitudes of reward and punishment respectively received by a person.

In all of the operant learning models, an action was the result of the application of two choice probabilities: choosing to act toward one potential partner or the other (target probability), and choosing to reward or punish the selected partner (reward probability). The participants begin with identical probabilities of providing a reward (0.995)[7] and of choosing an actor to act toward (0.5 toward each of the other actor). Each act was then chosen by first applying the target probability to the choice of exchange partner and then the reward probability to the choice of reward or punishment. In each subsequent turn, these probabilities are modified by the outcome of the prior turn. Behaviors that are rewarded have their probability of occurrence increased (while the alternative is decreased), and the probabilities of behaviors that are punished are decreased (while the alternative is increased).

The Agency models. In each of the goal seeking (agency) models, a two level control system was instituted. Two different models were considered that vary the strategies (outputs) used at each of the two levels. These are illustrated in Figure 4. In both models, the top level goal was to earn points. In one model (RPRP, discussed below) the outputs (strategies, and also goals for the lower level control system) varied between (a) earning points from potential exchange partner number one and (b) earning points from potential exchange partner number two. With either of these goals, the lower level control system also had binary outputs: (a) give points (reward), and (b) take points away (punish).

[7] This assumes, in agreement with observations, a strong propensity to give rewards rather than punishments. The exact figure can be changed, but as we shall see, there is a strong tendency in many of the models for this initial probability to degrade dramatically over time.

(Figure 4 About Here)

In the other model (RRPP), the top level outputs were (a) earn points from being rewarding (give points), and (b) earn points from being punishing (taking points away). With either of these two goals, the lower level control system in this model also had binary outputs: either act toward partner number one or act toward partner number two. Note that both models produce the same four behaviors, but with different priorities. In the first the priority is toward one partner or the other, in the second the priority is toward being rewarding or punishing. In each model, each behavior is tried (and retried) to bring about the goal (match perceptions to the standard). If the primary goal is not achieved, after trying and retrying all possible behaviors, the high-level control system alters its output thus instituting the secondary goal and all behaviors are tried again. If the secondary goal is not achieved, the process starts again with the primary goal. Because the top-level control system looks for reward, it will switch lower level goals (its output) even if the lower level system is satisfied. Thus, when the costs associated with achieving the reward outweigh the reward, the goal is changed.

In considering alternative agency models, two parameters are varied. The first is the strategy that is used to bring about the goal. The two different models just described (RPRP and RRPP) may be viewed as alternative strategies. The second parameter that is varied is the persistence in trying a particular behavior to bring about the match between perception and standard. Five different levels of persistence were examined consisting of 3, 5, 7, 9, and 11 tries before giving up one the present behavior as an effective technique and moving on to the next behavior in the sequence. The tries, however, were not fixed at these levels, but consisted of a random number of tries up to the specified level with a minimum of two tries. Thus for the level three, half the time the agent tried two times and half the time three tries.

The Outcomes

In order to assess the degree to which the activity of the participants becomes socially structured (as opposed, perhaps, to individually structured), we need to see that the activities are patterned, differentiated and correlated across participants. That is, in order to show the emergence of social structure I must show that, (1) actors in different positions engage in different activities or in different amounts of the same activities, (2) the amount of activity of an actor in one position is correlated with the amount of activity of an actor in another position, and (3) the correlations in activities are not the same across all pairs of positions.

The individual outcomes considered are (1) the proportion of exchanges by P1 and P2 that give rewards (across both potential exchange partners) and (2) the proportion of exchanges by P1 and P2 that are directed toward each other (toward *person* as opposed to toward *other*). At issue are variations in these outcomes between actor P1 and actor P2 (recall that O1 and O2 are symmetric with P1 and P2), and variations in these outcomes among the eight conditions in the distribution of reward and punishment power.

The group level outcome consists of the pattern of correlations among the individual behaviors. These correlations indicate patterns of emerging structure in the groups, as that structure may vary across the eight conditions of reward and punishment power.

Finally, all of the above outcomes will be compared across the different models that are simulated in order to learn about the effects of the basic assumptions concerning the nature of the individuals that comprise the interacting groups (i.e., forward looking versus backward looking models and variations of them).

Results

General Results

I begin with some general comparisons of the operant learning models with the agent models. Table 1 presents means and standard deviations for individual outcomes for five variations of operant learning models and for ten variations of identity theory (agent) models. I look first at the operant learning models.

(Table 1 About Here)

Operant Learning Models

With respect to the five operant learning models investigated, Table 1 shows that although actors P1 and P2 occupy structurally different positions, in terms of reward and punishment power, they do not differ in their behaviors of providing rewards to others or choosing to exchange with each other rather than their counterparts. Table 2 presents the correlations among the choices to reward and to exchange with person rather than other for three exemplar conditions. The full set of correlations is presented in the appendix. With respect to the correlations among the choices to reward and to exchange with person (P1 or P2) vs. other (O1 or O2), there are two different patterns. Each is illustrated in Table 2.

(Table 2 About Here)

For the activity of rewarding, reported in the first three rows of each table, there are either very weak (near zero) correlations among the members, as in the "seek improvement" model or the correlations are very strong (high 0.80s and 0.90s), as in the "neutral middle" model. This means that actors in one position are either providing rewards independently of the degree to which actors in other positions provide rewards (low correlations), or the actors within groups are all providing rewards to the same extent in each group (high correlations), perhaps independently responding to common stimuli. For the activity of exchanging with Person rather than Other, reported in the last three columns of each table, all the correlations are near zero, indicating each actor is making choices about exchange partners independent of any other actor's choices.

Finally, the correlations between the two activities across actors are also uniformly near zero in all the models. For example the degree to which P1 gives rewards is unrelated to the degree to which P1 also chooses to interact with person P2 (row 4 column 1 of the tables), being 0.01 or 0.02 in the models illustrated. These low correlations indicate that the extent of providing points (rewarding) by one actor has nothing to do with their own or others' selecting person or other as an exchange partner. Thus, I must conclude that no social structure in emerging in these groups, no sustained patterns of differentiated and interrelated activity.

Agent Models Based on Identity Theory

Considering the agent or identity theory models, Table 1 shows that actors P1 and P2 do differ from each other in the extent to which they reward others and choose each other rather than their counterparts. Uniformly, actor P1 uses less punishment than P2 (is more likely to reward) and is less likely to choose person than is P2.

With respect to the coordination of activities with others in the group, Table 2 (in conjunction with Table 1) shows one illustrative example of a pattern that is quite constant across all the agent models.[8] In this pattern, we see that actors P1 and O1 are in a coordinated "alliance," which actors P2 and O2 attempt to break up. To see this, I note that P1 and O1 vary their level of rewarding behavior together (r = 0.87) and the degree to which they choose each other (r = -0.97, indicating that as P1 chooses other, i.e. O1, O1 chooses person i.e., P1). P2 and O2 also vary their level of rewarding behavior together (r = 0.94). However, high levels of rewarding on the part of P1 and O1 are associated with low levels of rewarding on the part of P2 and O2 (r averaging about -.25). The degree to which P1 chooses P2, on the other hand (which is almost never as indicated in Table 1), is strongly related (r = 0.96) to the degree that P2 provides rewards (when P1 does not choose P2, P2 punishes P1). This pattern also shows in the strong positive correlation between P2's providing rewards and P2's choosing P1 (r = 0.73). When P2 chooses P1 and it is reciprocated, P2 provides rewards, otherwise P2 tends to punish more.

Overall, these results suggest that some degree of coordinated activity or rudimentary social structure has emerged for the agent (identity) models. Unlike the results for the operant models, we see that the amounts of each activity vary by structural position in the group, and are correlated across positions in the group.

Variations by Structural Conditions

Each of the models were run under each of the eight different structural conditions defined by the distribution of reward and punishment power (see Figure 3). I now discuss the results of analyses measuring the impact of the structural conditions on each of the outcomes discussed above.

[8] The full results for all models are shown in the Appendix.

Operant Learning Models.

Analysis of variance with a 2x2x2 design was used to estimate the impacts of the basic conditions of high vs. low punishment power, high vs. low reward power, consistent vs. inconsistent relationship between punishment and reward power, and their interactions on the behaviors of actors P1 and P2. Considering the 140 possible effects with the 5 models and 4 outcomes, only 10 were significant at the 0.05 level. Ten significant outcomes could easily have happened by chance ($p = 0.22$). In addition, the ten showed no pattern across models or behaviors. The average R^2 for the full set of results (see Table 3) was only 0.01. I must, therefore, conclude that the various conditions had no significant impact on the behaviors of actors P1 and P2 in any of the operant models.

Agent Models Based on Identity Theory

Analysis of variance of the agent models reveals a very different picture. For all of the behaviors of actors P1 and P2, for every agent model, there were strong, significant effects of power distribution, including the highest order three-way interaction of reward power, punishment power, and direction of punishment power. As reported in Table 3, the average R^2 for these results was 0.72, with half of them above 0.90. The fact that the three-way interaction is significant means that none of the effects of power distribution on the behaviors of P1 and P2 can adequately be described without taking into account all of the other effects. Thus, while the behaviors of the actors vary by the power distribution conditions, it does not appear to vary in any simple manner.

(Table 3 About Here)

Much of the complexity of the results, however, is only the consequence of the way we have been analyzing the data, and is not due to complexity of the model results. Examination of the means by condition revealed that most of the variation in the outcomes was the result of differences between conditions 1, 2, 3, and 5 and conditions 4, 6, 7, and 8 of Figure 3, and similarity within groupings. Some factor underlying these groupings appears to be an important determinant of the behavior of the actors. This factor, not readily apparent in the original 2x2x2 design, distinguishes between structures that are *attention-shifting* for actor P1 (conditions 1, 2, 3, and 5) and those that are *not* (conditions 4, 6, 7, and 8). [9] In the structures that are *not* attention-shifting for actor P1, P1 earns more from O1, *even when P2 punishes P1*, than could ever be earned by P1 from P2.[10] Thus, *P1 has no incentive to shift attention to P2* in these conditions. And, the results show that in these conditions P1 *never* chooses to interact with P2. While P1 can always earn most from O1, in the attention-shifting structures P1 can earn more from P2 than from-O1-while-being-punished-by-P2. Here, punishment by P2 motivates a shift in attention, or orientation, or goal by P1. This single binary factor, whether or not the network is attention-shifting for actor P1, is used to reanalyze the data.

The results of this reanalysis, also shown in Table 3, indicate that this single factor accounts for an average of 90.53% of the variance accounted for by the full 2x2x2 model. It is clearly this single factor which accounts for much of the apparent complexity of the results obtained in the full analysis. The fact that the attention-shifting structures occur in the cells they do (conditions 1, 2, 3, and 5) guarantees high-order interactions. But, it is

---

[9] Indeed, it was only after studying the pattern of results extensively, after the fact, that this factor came to be understood. Of course, afterwards, it is very obvious.

[10] This is calculated as the amount that O1 can reward P1 minus the amount that P2 can punish P1. This

not the high vs. low reward and punishment power *per* se that is producing the attention-shifting structure. This is clear because attention-shifting structures exist in all conditions: high or low reward power, high or low punishment power, punishment power aligned or not aligned with reward power. Rather it is particular relationships between the amounts that O1 can reward P1, that P2 can reward P1, and that P2 can punish P1 which define an attention-shifting structure, and this structure is important because P1 is an active agent that seeks the best reward possible.

Substantively, the analysis of the attention-shifting network factor shows that under the non-attention-shifting condition, P1 uniformly earns more while P2 earns less, both P1 and P2 pay less attention to each other, and P2 uniformly uses more punishment.[11] The non-attention-shifting structure also results in P1 using less punishment in the punishment strategy (RPRP), but has no impact on P1's use of punishment under the reward strategy (RRPP). Overall, both P1 and P2 are strongly influenced by the distribution of reward and punishment power, insofar as that distribution results in or does not result in an attention-shifting structure.

Variations in the Models

Before turning to consider some variations in the agency model, I should point out that variations in the learning models hinged on the definition of what was taken as reinforcement and punishment. The sole impact of this variation was to change the overall

---

amount is greater than the amount that P2 can reward P1.

[11] These results are not reported in a table, but correspond to the substantive results associated with the R^2 values reported in Table 3.

average amount of reward or punishment used by all actors in the group, as shown in Table 1. The seek reward and seek improvement models had a very restricted definition of reinforcement and the initial probability of providing reward (0.995) deteriorated very rapidly to a 50/50 choice with no correlations in the choice across members (from Table 2). In the other conditions, there was also deterioration, but not as rapidly, and not uniformly across groups. The result of this was high correlations in the giving of rewards as in some groups everyone gave more rewards and in some groups everyone gave fewer rewards (as the probability of giving rewards diminished). In none of the cases is this evidence of structure.

Turning now to variations in the agent/identity models, these variations occurred with respect to strategies used by the actors. It will be recalled that strategy was varied along two dimensions as shown in Figure 4. The first dimension concerned the ordering of behaviors tried. One kept a *focus on a prime exchange partner*, i.e., the other who could reward the most, and tried both rewarding and punishing behaviors before switching to the secondary exchange partner. This was the RPRP model. The other strategy tried to *maintain positive behavior* and switched exchange partners (from prime to secondary) before changing to negative behavior: the RRPP model. The second dimension of strategy pursued by agents concerned variation in the persistence with which any particular behavior was attempted before giving it up as ineffective.

What effects do variations in strategy have on the behaviors and patterns of interaction in these simulated groups? The results in Tables 1 and 2 show there is no simple answer, as the effects of variations in one dimension of strategy depend upon the other strategic dimension that is used, and some of the effects, especially in the correlations in Table 2, appear to be curvilinear. Worse than that, each of these variations in strategy interacts with variations in the reward and punishment power structure of the group (i.e., the attention-shifting character of that structure).

Table 4 shows the average levels of giving rewards and choosing person, for actors P1 and P2, by strategy and power structure. Table 5 presents analysis of variance results for these data. Looking at Table 4 we see that in the non-attention-shifting condition, actor P1 *never* exchanges with P2 (probability of exchanging with person is 0.00) and *always* rewards O1 (probability of rewarding is 1.00). In the attention-shifting condition, actor P1 has a very high probability of providing rewards using the RRPP strategy (0.98 to 1.00), but only a moderately high probability using the RPRP (0.66 to 0.83). This latter strategy results in much more punitive behavior (i.e., less rewarding), though the degree of punitiveness decreases as persistence increases. In addition, in the attention-shifting condition, actor P1 is much more likely to exchange with actor P2 (being close to 0.40 rather than 0.00). The effects of persistence on this likelihood differ according to whether P1 is using the RRPP strategy or the RPRP strategy. In the former, persistence increases the likelihood that P1 will choose P2 (rising from 0.33 to 0.42), while in the latter, persistence decreases the likelihood that P1 will choose P2 (decreasing from 0.43 to 0.37).

(Tables 4 and 5 about here)

Turning to a consideration of the behavior of actor P2, Table 4 shows that in the non-attention-shifting condition, P2 provides rewards between 55% and 66% of the time, with the lower number the result of using the RPRP strategy and the latter the result of using the RRPP strategy. These probabilities are not influenced by persistence. In the attention-shifting structure, P2 is more likely to provide rewards, again doing so more often in the

RRPP than the RPRP condition, and here we see that persistence has an effect (because P1 does pay some attention) with greater use of reward under higher levels of persistence.

With respect to choosing an exchange partner, in the non-attention-shifting condition, actor P2 chooses P1 about two-thirds of the time with little difference between using the RRPP or RPRP strategies, and with no variation by level of persistence. Again, the action comes in the attention-shifting condition. Here actor P2 is more likely to choose P1, and that choice is more likely to be made under conditions of more persistence (especially using the RPRP strategy). In sum, the level of giving rewards and of choosing and exchange partner are complex functions of the conditions in which the actors find themselves *and* the strategies they use to accomplish their goals.

Tables 6 and 7 present the results of a similar analysis of the correlations between the behaviors of the actors in the group. Once again, we see the complexity of the effects of conditions and strategies. In the non-attention-shifting conditions the correlations in the behavior between actors P1 and O1 and between P1 and P2 are uniformly zero because P1 and O1 *always exchange* with each other, and they *always reward* each other. In the attention-shifting structures, the exclusive relationship between P1 and O1 is broken up, and other patterns emerge. There is more coordination in the giving of rewards between P1 and P2 and between P1 and O1 when all use the RPRP strategy than when they use the RRPP strategy. The effect of persistence, however, is to lower that coordination between P1 and P2, but raise it between P1 and O1 in both conditions. On the other hand, that same persistence increases the coordination in giving rewards and choosing each other between P1 and O1. In general, strategies that increase the coordination between P1 and O1 (high persistence and RRPP strategies) decrease the coordination between P1 and P2.

(Table 6 and 7 About Here)

Discussion

Lying between the effects of social structure on the individual and the temporally subsequent impact of interacting individual on social structure is the nature of the individual and the kinds of interaction patterns that individuals of various natures might produce. Two fundamentally different views of the nature of the individual have existed in the sociological literature. The first, which has served as the model of the individual in the social exchange paradigm, is the individual as an actor who chooses one action or another on the basis of a history of reinforcement. The second, which has served as a model of the individual in identity theory and affect control theory (Smith-Lovin and Heise 1988), is the individual as an actor/agent who chooses actions in order to accomplish an internally held goal. Simulations of interacting actors based on these models were developed to explore their implications. Simulations were used because the complexity and nonlinearities of a social system with several actors acting and reacting to each other over time prevents an analytic, deductive exploration. The results of the simulations led to some strong conclusions, at least for the settings involved in the present study.

First, *none* of the operant learning based models of actors interacting led to consistent, structured patterns of interaction, but *all* of the agent based (goal oriented) models did lead to such consistent, patterned structures of behavior and interaction. Second, the behaviors that evolved in the operant learning based models were *not* influenced by the structural conditions (relative levels of reward and punishment power) of the settings explored, while behaviors that evolved in the agent based models *were* influenced, some-times quite strongly, by the structural conditions. Third, the behavior of actors in structurally different positions within a group was not differentiated in the operant learning based models, but was quite strongly differentiated in the agent based models. For this reason, individual agency may be the *sine qua non* for the emergence of social structure.

Let me first discuss the interactions based on the operant learning model. As we saw, no patterns of interaction emerged in the operant learning groups,[12] yet we know that operant learning theory is well tested and has worked successfully in laboratory situations to account for many patterns of behavior. What happened in this context? Perhaps, operant learning theory works well only when an experimenter is able to control the situation including *all* rewards and punishments that a subject might be given. In that case, the subject receives a consistent, patterned set of rewards and punishments and learns to engage in the behavior sought or to make the discrimination desired. In the present simulations, however, no one is in charge. There is not just one behavior that is being learned, there is not just one source of reward or punishment, and the sources of reward and punishment don't have particular goals in mind. When considered in this fashion, it is no wonder that stable patterns of interaction do not emerge. This is not to say that in other structures, with other patterns of reward and punishment power that such patterns would not emerge, but it certainly raises the question. Without preexisting patterns in the actual use of reward and punishment, patterns in the behavior being rewarded and/or punished simply do not emerge.[13]

[12] This was true of all forms of the basic operant model and the satisfaction-balance model.

[13] An alternative to structuring the rewards and punishments for each actor (by an experimenter) would be

Turning now to the models based on agency, we have seen that all the agency models led to structured interaction, but different patterns of interaction developed and become stable under different structural conditions. For the conditions studied in the present simulations, the largest factor was whether or not the structure had what I termed an "attention-shifting" character or not. This characteristic of the structure involves the relative reward and punishment power of all four actors. In all of the conditions, Actor P1 could get a bigger reward from O1 than from P2, and was therefore initially oriented toward O1. Both actors P2 and O1 could get a bigger reward from P1 than from O2 and were initially oriented toward P1. The question, then, is whether P2 has sufficient reward and punishment power to shift P1's attention away from O1 toward P2. If O1's rewards to P1 are strong enough that punishment from P2 does not cause P1 to shift attention to exchanges with P2, we do not have an attention-shifting structure. In the present simulations, this condition accounted for most of the variance in all of the outcomes for all of the agent models.

There is a general principle here having to do with the nature of reward and punishment power in agent models of interaction. *It does not matter how much power one agent has relative to another, if the power is not sufficient to produce an attention-shift in that other, there may as well be no power*. What this study reveals is that there is a funda-

---

to have each simulated actor focus selectively on rewards and punishments in the situation. This alternative was simulated for both the Neutral Middle and Satisfaction-Balance models by having the actors attend only to the rewards and punishments emanating from the target actor of their last act. This yielded results that began to show signs of the emergence of structure (Burke and Gray 1997). However, such focus constitutes the beginnings of an internal standard as in the agent/identity model.

mental discontinuity in power magnitudes, which depends upon the nature of the social structure that defines power distributions. Power levels below the attention-shifting threshold are effectively zero. Power is the ability to disrupt the controlled perceptions of others in such a way as to cause their “corrective” behavior to be in line with one’s own goals. Without being able to disturb the perceptions that the other is controlling, there is no power. Thus, power is the ability to influence how another agent *pursues its own goals*.

When looking at the behavior of P1, who is initially oriented toward O1 in the present structures, if P2’s power is below the attention-shift threshold, the interaction between P1 and O1 is stable. If P2’s power is above the attention-shift threshold, P1 will shift orientation from O1 to P2, so there is some initial instability. The question then becomes whether O1 has sufficient reward and punishment power to be above an attention-shifting threshold with respect to P1. That is, can O1 recapture P1’s attention. If not, then after the initial instability of P1 shifting attention from O1 to P2, the interaction patterns stabilize with P1 and P2 involved in exchanges. If, on the other hand, O1 does have enough power to be above the attention-shift level, we have an unstable pattern of interaction emerging in which P1 interacts first with O1 until P2 captures P1’s attention then P1 interacts with P2 until O1 captures P1’s attention. P1 will shift back and forth between O1 and P2 in an attempt to gain first one goal and then another, unable to fully satisfy either. In the present experimental conditions, O1 has sufficient power in all of the conditions to bring about an attention-shift in P1.

Again, there is a general principle here concerning the stability of interaction in networks of the present general form. *From the point of view of the control system involved* *in identity theory, completely stable interaction structures emerge only when all actors are achieving the goals they set for themselves*, that is, when their perceptions match their standards. Under this condition, the error signal is zero and no further changes in behavior are needed. However, when several actors are present in a situation, and the outputs of one actor create disturbances in the perceptions of others, it is possible that such disturbances cannot be countered without disturbing the perceptions of other actors, as in the present instance. In this situation, where the behaviors needed by one actor to achieve congruency create disturbances to the congruency between perceptions and standards of others, one can speak of conflict between the actors (Powers 1973).

When such conflicts arise, if persons cannot leave the situation (Swann 1987), they will need to adjust their *standards*, not simply their behavior, in order to achieve congruency. Under some conditions, the adjusted goals (and behaviors) will lead to congruency between the perceptions and standards of all the actors, but under other conditions, there is no arrangement of goals (and behaviors) that is mutually satisfactory. Thus, I distinguish among three conditions. The first is where either there is congruency between goals and perceptions or such can be achieved with slight modifications in behavior. The second is where congruency, though not currently present, can be obtained by modifying both goals and behavior. The third is where congruency cannot be obtained because no combination of goals and behavior can be found that is mutually satisfactory.

In the present simulations we have conditions two and three above. There is initial instability because congruency cannot be obtained between the *primary* goals and perceptions of all the actors at the same time. In the non-attention shifting groups P1 and O1 are oriented toward each other and P2 and O2 must settle for each other after an initial discovery that they do not have the power to change the orientation of P1 or O1. In the attention-shifting groups, no final stability can emerge, since P2 and O1 each has the power to change the orientation of P1 (their primary goal of an exchange partner).

Finally, I raise the question of the relationship between social structure and different agent strategies. The consequences of the different strategies that were modeled were seen to be clearly contingent upon the social structure of reward and punishment power distributions within which they were employed. This leads to the final principle emerging from this research. *The meanings or consequences of particular behaviors within a social structure depend very strongly upon that structure, and consequently yield different patterns of interaction under different conditions.* What has not been modeled, but is a clear direction for future research, is to make the distribution of reward and punishment power in part dependent upon the patterns of interaction that emerge, thus making social structure more realistically and fully dependent upon emerging patterns of interaction. In the present laboratory conditions, the actors did not gain points by taking them away from others, not did they loose points by giving them to others. Adding this feature may help close the loop on understanding the micro-macro problem of the relationship between the individual and society.

PostScript

The interaction patterns of “real” people under the different conditions of reward and punishment power have not been considered in this paper in order not to distract from comparisons of the implications of the theoretical models investigated. To see how people actually behave under these conditions the reader is encouraged to go to the original report by Molm (1989). Nevertheless, a couple of words are in order. Neither the agent models nor the operant learning models behave like real persons under these conditions, though in many respects the agent models are much closer. The biggest difference, of course, is that real people are not confined to the limited goal set that characterize the agents being modeled, and people have affective as well as cognitive goals to contend with. For example, Molm points out that people in the real groups often used punishment as a retaliation for being punished. In addition, the agent models are single minded in the pursuit of their goals and proceed systematically, never missing nor misinterpreting information they are given, consequently the attention-shifting character of some of the networks fails to have the impact for real persons that it has for these agents. In order to model the way people act in the present networks, additional goals would need to be built into the agents, and perhaps mechanisms for being distracted or failing to take full account of information provided. It is difficult to imagine, however, modifications of the operant learning models that would yield realistic results.

References

Burgess, Robert L and Jr. Bushell, Don, Editors. 1969. *Behavioral Sociology*. New York: Columbia University Press.

Burke, Peter J. 1991. "Identity Processes and Social Stress." *American Sociological Review* 56:836-49.

______. 1997. "An Identity Model for Network Exchange." *American Sociological Review* 62:134-50.

Burke, Peter J. and Alicia D. Cast. 1997. "Stability and Change in the Gender Identities of Newly Married Couples." *Social Psychology Quarterly* 60:277-90.

Burke, Peter J. and Louis N. Gray. 1997. "Where Forward-Looking and Backward-Looking Models Meet." Mathematical Sociology Section Session at the Annual Meetings of the American Sociological Association. Toronto, August.

Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, MA: Belknap Press of Harvard University Press.

Cook, Karen S. and Richard M. Emerson. 1978. "Power, Equity and Commitment in Exchange Networks." *American Sociological Review* 43:721-39.

Cook, Karen S. and Toshio Yamagishi. 1992. "Power in Exchange Networks: A Power-Dependency Formulation." *Social Networks* 14.

Emerson, Richard M. 1962. "Power-Dependence Relations." *American Sociological Review* 27:31-40.

______. 1969. "Operant Psychology and Exchange Theory." Pp. 379-405 in *Behavioral Sociology*, edited by Robert L. Burgess and Don Bushell, Jr. New York: Columbia University Press.

Estes, William K. 1957. "Of Models and Men." *American Psychologist* 12:609-17.

Freese, Lee and Peter J. Burke. 1994. "Persons, Identities, and Social Interaction." Pp. 1-24 in *Advances in Group Processes*, vol. 11, edited by Barry Markovsky, Karen Heimer and Jodi O'Brien. Greenwich, CT: JAI Press, Inc.

Gray, Louis N. and Irving Tallman. 1984. "A Satisfaction Balance Model of Decision Making and Choice Behavior." *Social Psychology Quarterly* 47:146-59.

______. 1986. "Predicting Choices in Asymptotic Decisions: A Comparison of Two Models." *Social Psychology Quarterly* 49:201-06.

______. 1987. "Theories of Choice: Contingent Reward and Punishment Applications." *Social Psychology Quarterly* 50:16-25.

Herrnstein, Richard. 1970. "On the Law of Effect." *Journal of the Experimental Analysis of Behavior* 17:243-66.

Homans, George C. 1961. *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace & World, Inc.

Molm, Linda D. 1989. "Punishment Power: A Balancing Process in Power-Dependence Relations." *American Journal of Sociology* 94:1392-418.

Nowak, Andrzej and Maciej Lewenstein. 1994. "Dynamical Systems: A Tool for Social Psychology?" Pp. 17-53 in *Dynamical Systems in Social Psychology*, edited by Robin R. Vallacher and Andrzej Nowak. New York: Academic Press, Inc.

Powers, William T. 1973. *Behavior: The Control of Perception*. Chicago: Aldine Publishing Co..

Skinner, P2.F. 1974. *About Behaviorism*. New York: Alfred A. Knopf.

Smith-Lovin, Lynn and David R. Heise, eds. 1988. *Analyzing Social Interaction: Advances in Affect Control Theory*. New York: Gordon and Breach Science Publishers.

Swann, William P2., Jr. 1987. "Identity Negotiation: Where Two Roads Meet." Journal of Personality and Social Psychology 53:1038-51.

Thibaut, John W. and Harold H. Kelley. 1959. *The Social Psychology of Groups*. New York: John Wiley & Sons, Inc.

Turner, Ralph. 1988. "Personality in Society: Social Psychology's Contribution to Sociology." Social Psychology Quarterly 51:1-10.

Table 1. Means and Standard Deviations for Individual Outcomes of Probability of Providing a Reward (Points), and Probability of Exchanging with Person Rather Than Other* for Actors P1 and P2 (N=2000/cell).**

	Pr. Rewarding					Pr. Choose Person				
	Actor P1			Actor P2		Actor P1			Actor P2	
	Mean	sd		Mean	sd	Mean	sd		Mean	sd
Operant Learning Models										
Avoid Punishment	0.78	(0.20)		0.77	(0.20)	0.51	(0.24)		0.51	(0.24)
Seek Reward	0.53	(0.04)		0.53	(0.04)	0.50	(0.04)		0.50	(0.04)
Seek Improvement	0.53	(0.05)		0.53	(0.04)	0.50	(0.04)		0.50	(0.05)
Neutral Middle	0.89	(0.16)		0.89	(0.15)	0.51	(0.26)		0.51	(0.26)
SB Model	0.92	(0.14)		0.92	(0.14)	0.52	(0.29)		0.51	(0.29)
Agent Models with RRPP Strategy										
3 Tries	0.99	(0.01)	**	0.77	(0.09)	0.16	(0.16)	**	0.68	(0.04)
5 Tries	1.00	(0.01)	**	0.79	(0.10)	0.19	(0.19)	**	0.71	(0.04)
7 Tries	1.00	(0.01)	**	0.80	(0.11)	0.20	(0.20)	**	0.71	(0.05)
9 Tries	1.00	(0.01)	**	0.80	(0.12)	0.20	(0.21)	**	0.70	(0.05)
11 Tries	1.00	(0.01)	**	0.80	(0.13)	0.21	(0.21)	**	0.70	(0.06)
Agent Models with RPRP Strategy										
3 Tries	0.83	(0.17)	**	0.57	(0.02)	0.21	(0.21)	**	0.61	(0.03)
5 Tries	0.87	(0.13)	**	0.59	(0.05)	0.20	(0.20)	**	0.64	(0.06)
7 Tries	0.89	(0.11)	**	0.61	(0.07)	0.19	(0.19)	**	0.66	(0.08)
9 Tries	0.91	(0.10)	**	0.62	(0.08)	0.19	(0.19)	**	0.67	(0.09)
11 Tries	0.92	(0.09)	**	0.63	(0.10)	0.18	(0.19)	**	0.68	(0.10)

* Other for P1 is O1, other for P2 is O2.

** Significant differences between Actors P1 and P2 in outcomes

Table 2. Correlations of the Activities of Participants
P1, P2, O1, and O2 in Selected Models.*

Illustrative Operant Learning Models								
Seek Improvement								
		Activity						
		Providing Rewards				Choosing Person as Target		
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	0.08						
	O1	0.02	0.02					
	O2	-0.01	0.03	0.02				
Choose Person	P1	0.01	0.02	0.02	0.03			
	P2	-0.02	-0.02	0.03	0.01	0.01		
	O1	-0.02	0.00	0.03	-0.05	0.00	0.01	
	O2	0.01	-0.01	-0.01	-0.02	-0.01	0.04	0.02
Neutral Middle								
		Activity						
		Providing Rewards				Choosing Person as Target		
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	0.95*						
	O1	0.96*	0.93*					
	O2	0.93*	0.96*	0.95*				
Choose Person	P1	0.02	0.03	0.02	0.03			
	P2	0.02	0.02	0.02	0.02	0.07		
	O1	0.01	0.00	0.01	0.00	-0.07	0.05	
	O2	0.02	0.02	0.02	0.02	0.05	-0.02	0.00
Illustrative Agent Model								
7 Tries with RRPP strategy								
		Activity						
		Providing Rewards				Choosing Person as Target		
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	-0.25*						
	O1	0.87*	-0.25					
	O2	-0.27*	0.94	-0.25				
Choose Person	P1	-0.28*	0.96*	-0.26*	0.96*			
	P2	-0.07	0.73*	-0.10	0.63*	0.68*		
	O1	0.29*	-0.95*	0.29*	-0.96*	-0.97*	-0.70*	
	O2	0.08	-0.63*	0.04	-0.74*	-0.70*	-0.33*	0.67*

+ Full results are contained in the appendix.
* $p \leq .05$

Table 3. Variance Accounted For (R^2) By Reward and Punishment Power Conditions in Full Model and By Attention Shift Code Alone for All Models.

	R^2 for full 2x2x2 model				R^2 for Attention Shift			
	Reward Others		Choose Person		Reward Others		Choose Person	
	P1	P2	P1	P2	P1	P2	P1	P2
Operant Learning Models								
Avoid Punishment	.02	.02	.00	.00	.00	.00	.00	.00
Seek Reward	.01	.00	.00	.00	.00	.00	.00	.00
Seek Improvement	.01	.00	.01	.00	.00	.00	.00	.00
Neutral Middle	.01	.01	.01	.00	.00	.00	.00	.00
SB Model	.01	.01	.01	.00	.00	.00	.00	.00
Agent Models with RRPP Strategy								
3 Tries	.37	.92	.97	.29	.37	.92	.97	.28
5 Tries	.19	.94	.97	.54	.15	.94	.97	.54
7 Tries	.15	.93	.96	.46	.08	.93	.96	.45
9 Tries	.14	.93	.96	.48	.07	.93	.96	.47
11 Tries	.09	.92	.96	.37	.05	.92	.96	.36
Agent Models with RPRP Strategy								
3 Tries	.99	.57	.99	.18	.99	.55	.99	.07
5 Tries	.97	.78	.98	.53	.95	.77	.98	.33
7 Tries	.94	.80	.96	.59	.90	.79	.96	.42
9 Tries	.93	.80	.94	.61	.86	.79	.94	.44
11 Tries	.90	.81	.94	.57	.83	.80	.93	.44

Table 4. Average Probability of Actors P1 and P2 Providing Rewards (Points) and Choosing Person By Level of Persistence, Reward Strategy, and Attention-Shifting Structure

Structure	Strategy	Persistence (number of tries) 3	5	7	9	11	Total
Probability that Actor P1 Rewards							
Non-Shifting	RRPP	1.00	1.00	1.00	1.00	1.00	1.00
	RPRP	1.00	1.00	1.00	1.00	1.00	1.00
	total	1.00	1.00	1.00	1.00	1.00	1.00
Shifting	RRPP	0.98	0.99	1.00	1.00	1.00	0.99
	RPRP	0.66	0.74	0.79	0.82	0.83	0.77
	total	0.82	0.87	0.89	0.91	0.91	0.88
Probability that Actor P2 Rewards							
Non-Shifting	RRPP	0.66	0.66	0.66	0.66	0.66	0.66
	RPRP	0.56	0.55	0.54	0.55	0.54	0.55
	total	0.61	0.61	0.60	0.60	0.60	0.61
Shifting	RRPP	0.85	0.89	0.90	0.91	0.92	0.89
	RPRP	0.59	0.64	0.67	0.70	0.71	0.66
	total	0.72	0.76	0.79	0.81	0.82	0.78
Probability that Actor P1 Chooses Actor P2							
Non-Shifting	RRPP	0.00	0.00	0.00	0.00	0.00	0.00
	RPRP	0.00	0.00	0.00	0.00	0.00	0.00
	total	0.00	0.00	0.00	0.00	0.00	0.00
Shifting	RRPP	0.33	0.38	0.40	0.41	0.42	0.39
	RPRP	0.43	0.39	0.38	0.37	0.37	0.39
	total	0.38	0.39	0.39	0.39	0.39	0.39
Probability that Actor P2 Chooses Actor P1							
Non-Shifting	RRPP	0.66	0.66	0.66	0.66	0.65	0.66
	RPRP	0.60	0.61	0.61	0.61	0.62	0.61
	total	0.63	0.63	0.63	0.63	0.63	0.63
Shifting	RRPP	0.70	0.73	0.73	0.73	0.73	0.72
	RPRP	0.61	0.68	0.71	0.73	0.75	0.70
	total	0.65	0.70	0.72	0.73	0.74	0.71

Table 5. Significance Tests for the Effects of Agent Strategies and Attention-Shifting Structures from the Analysis of Variance in the Probabilities of Actors Providing Rewards and Choosing Person as an Exchange Partner.

		Analysis of the probability of providing rewards					
		Actor P1			Actor P2		
Source	df	MS	F	p <	MS	F	p <
Model[+]	19	11.49	20473.73	0.00	17.39	16328.17	0.00
MAX	4	1.40	2494.90	0.00	1.36	1273.43	0.00
PUN	1	63.48	113150.46	0.00	150.87	141636.73	0.00
Ashift	1	71.71	127818.16	0.00	149.48	140332.51	0.00
MAX*PUN	4	1.05	1865.32	0.00	0.07	64.25	0.00
MAX*Ashift	4	1.40	2494.90	0.00	1.77	1665.14	0.00
PUN*Ashift	1	63.48	113150.46	0.00	16.72	15701.48	0.00
MAX*PUN*Ashift	4	1.05	1865.32	0.00	0.15	138.32	0.00
Residual	19980	0.00			0.00		
R-square				0.95			0.94

		Analysis of the probability of choosing person					
		Actor P1			Actor P2		
Source	df	MS	F	p <	MS	F	p <
Model[+]	19	40.16	27769.52	0.00	2.66	1037.67	0.00
MAX	4	0.04	30.60	0.00	1.36	532.12	0.00
PUN	1	0.00	0.78	0.38	6.66	2597.30	0.00
Ashift	1	755.33	522278.33	0.00	30.51	11902.11	0.00
MAX*PUN	4	0.92	637.03	0.00	0.61	237.02	0.00
MAX*Ashift	4	0.04	30.60	0.00	0.95	372.13	0.00
PUN*Ashift	1	0.00	0.78	0.38	0.68	263.81	0.00
MAX*PUN*Ashift	4	0.92	637.03	0.00	0.25	96.88	0.00
Residual	19980	0.00			0.00		
R-square				0.96			0.50

[+] MAX = effects of maximum number of tries before changing behavior (persistence),
PUN = effects of RRPP vs. RPRP strategy,
Ashift = effects of attention-shifting vs. non-attention-shifting structures.

Table 6. Average Correlations Between Actors of (a) Their Probabilities of Providing Rewards and (b) Their Probabilities of Choosing Person By Level of Persistence, Reward Strategy and Attention-Shifting Structure.

Correlations in providing rewards.

		Persistence (number of tries)					
Structure	Strategy	3	5	7	9	11	Total
Correlations Between Actors P1 and P2							
Non-Shifting	RRPP	0.00	0.00	0.00	0.00	0.00	0.00
	RPRP	0.00	0.00	0.00	0.00	0.00	0.00
	Total	0.00	0.00	0.00	0.00	0.00	0.00
Shifting	RRPP	0.31	0.20	0.15	0.13	0.15	0.19
	RPRP	0.52	0.54	0.49	0.40	0.43	0.48
	Total	0.42	0.37	0.32	0.26	0.29	0.33
Correlations Between Actors P1 and O1							
Non-Shifting	RRPP	0.00	0.00	0.00	0.00	0.00	0.00
	RPRP	0.00	0.00	0.00	0.00	0.00	0.00
	Total	0.00	0.00	0.00	0.00	0.00	0.00
Shifting	RRPP	0.74	0.80	0.84	0.85	0.85	0.82
	RPRP	0.45	0.52	0.55	0.50	0.55	0.52
	Total	0.59	0.66	0.70	0.67	0.70	0.67
Correlations Between Actors P2 and O2							
Non-Shifting	RRPP	0.07	0.07	0.20	0.28	0.26	0.18
	RPRP	0.20	0.06	0.01	0.01	0.02	0.06
	Total	0.14	0.07	0.10	0.14	0.14	0.12
Shifting	RRPP	-0.19	-0.11	-0.13	-0.16	-0.14	-0.14
	RPRP	-0.06	-0.13	-0.20	-0.31	-0.25	-0.19
	Total	-0.12	-0.12	-0.16	-0.23	-0.19	-0.17

Table 6. (cont.) Average Correlations Between Actors of (a) Their Probabilities of Providing Rewards and (b) Their Probabilities of Choosing Person By Level of Persistence, Reward Strategy and Attention-Shifting Structure.

Correlations in choosing person.

Structure	Strategy	Persistence (number of tries) 3	5	7	9	11	Total
Correlations Between Actors P1 and P2							
Non-Shifting	RRPP	0.00	0.00	0.00	0.00	0.00	0.00
	RPRP	0.00	0.00	0.00	0.00	0.00	0.00
	Total	0.00	0.00	0.00	0.00	0.00	0.00
Shifting	RRPP	-0.05	0.04	0.08	0.08	0.12	0.05
	RPRP	-0.02	0.04	0.08	0.22	0.13	0.09
	Total	-0.04	0.04	0.08	0.15	0.12	0.07
Correlations Between Actors P1 and O1							
Non-Shifting	RRPP	0.00	0.00	0.00	0.00	0.00	0.00
	RPRP	0.00	0.00	0.00	0.00	0.00	0.00
	Total	0.00	0.00	0.00	0.00	0.00	0.00
Shifting	RRPP	-0.06	-0.05	0.02	0.07	0.17	0.03
	RPRP	0.42	0.56	0.64	0.71	0.69	0.60
	Total	0.18	0.26	0.33	0.39	0.43	0.32
Correlations Between Actors P2 and O2							
Non-Shifting	RRPP	0.34	0.43	0.37	0.43	0.40	0.39
	RPRP	-0.61	-0.19	-0.03	0.17	0.16	-0.10
	Total	-0.14	0.12	0.17	0.30	0.28	0.15
Shifting	RRPP	-0.08	0.12	0.20	0.18	0.25	0.13
	RPRP	0.04	-0.01	-0.01	-0.02	0.01	0.01
	Total	-0.02	0.07	0.09	0.10	0.13	0.07

Table 7. Significance Tests for the Effects of Agent Strategies and Attention-Shifting Structures from the Analysis of Variance in the Correlations Between Actors of Their Probabilities of Providing Rewards and Their Probabilities of Choosing Person.

		Analysis of Correlations in Providing Rewards Between								
		Actors P1 & P2			Actors P1 & O1			Actors P2 & O2		
Source	df	MS	F	p <	MS	F	p <	MS	F	p <
Total	79									
Model+	19	0.16	45.26	0.00	0.53	411.95	0.00	0.12	14.18	0.00
MAX	4	0.01	2.34	0.07	0.00	3.73	0.01	0.04	4.24	0.00
PUN	1	0.44	121.10	0.00	0.52	402.08	0.00	0.28	32.22	0.00
Ashift	1	2.14	596.31	0.00	8.96	6978.75	0.00	1.39	162.46	0.00
MAX*PUN	4	0.00	0.34	0.85	0.00	1.79	0.14	0.08	9.18	0.00
MAX*Ashift	4	0.01	2.34	0.07	0.00	3.73	0.01	0.03	3.38	0.01
PUN*Ashift	1	0.44	121.10	0.00	0.52	402.08	0.00	0.02	2.09	0.15
MAX*PUN*Ashift	4	0.00	0.34	0.85	0.00	1.79	0.14	0.01	1.35	0.26
Residual	60	0.00			0.00			0.01		
R-square				0.93			0.99			0.82

		Analysis of Correlations in Choosing Person Between								
		Actors P1 & P2			Actors P1 & O1			Actors P2 & O2		
Source	df	MS	F	p <	MS	F	p <	MS	F	p <
Total	79									
Model+	19	0.02	2.29	0.01	0.29	103.41	0.00	0.23	6.20	0.00
MAX	4	0.01	0.61	0.66	0.04	12.94	0.00	0.16	4.24	0.00
PUN	1	0.00	0.01	0.92	1.65	596.75	0.00	2.09	55.43	0.00
Ashift	1	0.24	25.92	0.00	1.81	653.45	0.00	0.11	2.93	0.09
MAX*PUN	4	0.01	1.58	0.19	0.00	1.79	0.14	0.04	1.07	0.38
MAX*Ashift	4	0.01	0.61	0.66	0.04	12.94	0.00	0.10	2.77	0.03
PUN*Ashift	1	0.00	0.01	0.92	1.65	596.75	0.00	0.47	12.57	0.00
MAX*PUN*Ashift	4	0.01	1.58	0.19	0.00	1.79	0.14	0.14	3.63	0.01
Residual	60	0.01			0.00			0.04		
R-square				0.42			0.97			0.66

+ MAX = effects of maximum number tries before changing behavior (persistence),
PUN = effects of RRPP vs. RPRP strategy,
Ashift = effects of attention-shifting vs. non-attention-shifting structures.

Appendix: The Full Results Illustrated in Table 2.
Correlations of the Activities of Participants P1, P2, O1, and O2.

		Activity						
		Providing Rewards				Choosing Person as Target		
Operant Learning Models								
Avoid Punishment								
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	0.93						
	O1	0.94	0.91					
	O2	0.91	0.93	0.93				
Choose Person	P1	0.00	0.01	0.00	0.00			
	P2	0.03	0.03	0.03	0.04	0.02		
	O1	0.02	0.01	0.02	0.02	0.03	0.01	
	O2	0.03	0.03	0.04	0.03	0.04	0.02	-0.05
Seek Reward								
Reward	P2	0.13						
	O1	0.16	0.02					
	O2	-0.01	0.07	0.14				
Choose Person	P1	0.01	0.10	-0.09	0.00			
	P2	0.11	0.01	0.00	-0.10	0.07		
	O1	0.09	0.01	0.00	-0.05	-0.03	-0.04	
	O2	0.00	0.07	-0.08	-0.01	0.02	-0.02	0.01
Seek Improvement								
Reward	P2	0.08						
	O1	0.02	0.02					
	O2	-0.01	0.03	0.02				
Choose Person	P1	0.01	0.02	0.02	0.03			
	P2	-0.02	-0.02	0.03	0.01	0.01		
	O1	-0.02	0.00	0.03	-0.05	0.00	0.01	
	O2	0.01	-0.01	-0.01	-0.02	-0.01	0.04	0.02
Neutral Middle								
Reward	P2	0.95						
	O1	0.96	0.93					
	O2	0.93	0.96	0.95				
Choose Person	P1	0.02	0.03	0.02	0.03			
	P2	0.02	0.02	0.02	0.02	0.07		
	O1	0.01	0.00	0.01	0.00	-0.07	0.05	
	O2	0.02	0.02	0.02	0.02	0.05	-0.02	0.00
SB Model								
Reward	P2	0.88						
	O1	0.91	0.84					
	O2	0.82	0.90	0.88				
Choose Person	P1	0.02	0.07	0.01	0.06			
	P2	-0.01	0.00	0.00	0.00	0.09		
	O1	0.02	-0.02	0.00	-0.03	-0.24	0.03	
	O2	0.04	0.03	0.04	0.02	0.00	-0.03	0.11

Appendix. (cont.) Correlations of the Activities of Participants P1, P2, O1, and O2.

		Activity						
		Providing Rewards				Choosing Person as Target		
Agent Models with RRPP Strategy								
3 Tries								
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	-0.50						
	O1	0.84	-0.55					
	O2	-0.57	0.90	-0.55				
Choose Person	P1	-0.60	0.93	-0.59	0.94			
	P2	-0.06	0.70	-0.15	0.50	0.54		
	O1	0.57	-0.94	0.60	-0.93	-0.97	-0.55	
	O2	0.25	-0.48	0.11	-0.68	-0.55	-0.33	0.51
5 Tries								
Reward	P2	-0.35						
	O1	0.84	-0.35					
	O2	-0.36	0.94	-0.32				
Choose Person	P1	-0.39	0.96	-0.36	0.96			
	P2	-0.12	0.79	-0.17	0.68	0.71		
	O1	0.39	-0.96	0.39	-0.96	-0.97	-0.74	
	O2	0.13	-0.61	0.05	-0.73	-0.68	-0.39	0.66
7 Tries								
Reward	P2	-0.25						
	O1	0.87	-0.25					
	O2	-0.27	0.94	-0.25				
Choose Person	P1	-0.28	0.96	-0.26	0.96			
	P2	-0.07	0.73	-0.10	0.63	0.68		
	O1	0.29	-0.95	0.29	-0.96	-0.97	-0.70	
	O2	0.08	-0.63	0.04	-0.74	-0.70	-0.33	0.67
9 Tries								
Reward	P2	-0.24						
	O1	0.86	-0.25					
	O2	-0.23	0.93	-0.23				
Choose Person	P1	-0.26	0.95	-0.27	0.95			
	P2	-0.07	0.71	-0.07	0.63	0.67		
	O1	0.25	-0.94	0.26	-0.95	-0.96	-0.69	
	O2	0.04	-0.59	0.04	-0.68	-0.66	-0.29	0.64
11 Tries								
Reward	P2	-0.19						
	O1	0.84	-0.19					
	O2	-0.20	0.93	-0.19				
Choose Person	P1	-0.20	0.95	-0.21	0.94			
	P2	0.00	0.65	-0.01	0.58	0.62		
	O1	0.23	-0.94	0.21	-0.95	-0.95	-0.64	
	O2	0.04	-0.59	0.06	-0.67	-0.66	-0.22	0.64

Appendix. (cont.) Correlations of the Activities of Participants P1, P2, O1, and O2.

		Activity						
		Providing Rewards				Choosing Person as Target		
Agent Models with RPRP Strategy								
3 Tries								
		P1	P2	O1	O2	P1	P2	O1
Reward	P2	-0.70						
	O1	0.99	-0.73					
	O2	-0.72	0.56	-0.70				
Choose Person	P1	-0.99	0.74	-0.99	0.72			
	P2	-0.20	0.61	-0.24	0.27	0.25		
	O1	0.99	-0.72	0.99	-0.74	-0.99	-0.23	
	O2	0.27	-0.28	0.23	-0.63	-0.26	-0.35	0.28
5 Tries								
Reward	P2	-0.80						
	O1	0.98	-0.83					
	O2	-0.82	0.76	-0.79				
Choose Person	P1	-0.97	0.87	-0.98	0.83			
	P2	-0.44	0.74	-0.49	0.56	0.56		
	O1	0.98	-0.83	0.97	-0.86	-0.96	-0.52	
	O2	0.48	-0.55	0.44	-0.74	-0.53	-0.56	0.55
7 Tries								
Reward	P2	-0.78						
	O1	0.98	-0.81					
	O2	-0.81	0.78	-0.78				
Choose Person	P1	-0.95	0.88	-0.96	0.83			
	P2	-0.45	0.76	-0.50	0.61	0.61		
	O1	0.96	-0.83	0.95	-0.89	-0.93	-0.58	
	O2	0.48	-0.58	0.44	-0.75	-0.55	-0.59	0.61
9 Tries								
Reward	P2	-0.76						
	O1	0.96	-0.80					
	O2	-0.80	0.76	-0.77				
Choose Person	P1	-0.92	0.89	-0.94	0.82			
	P2	-0.43	0.76	-0.49	0.58	0.64		
	O1	0.94	-0.81	0.92	-0.89	-0.91	-0.57	
	O2	0.49	-0.60	0.44	-0.75	-0.58	-0.56	0.64
11 Tries								
Reward	P2	-0.74						
	O1	0.96	-0.77					
	O2	-0.77	0.77	-0.73				
Choose Person	P1	-0.90	0.89	-0.92	0.81			
	P2	-0.41	0.75	-0.46	0.60	0.64		
	O1	0.92	-0.82	0.90	-0.88	-0.90	-0.59	
	O2	0.45	-0.58	0.40	-0.76	-0.58	-0.55	0.64