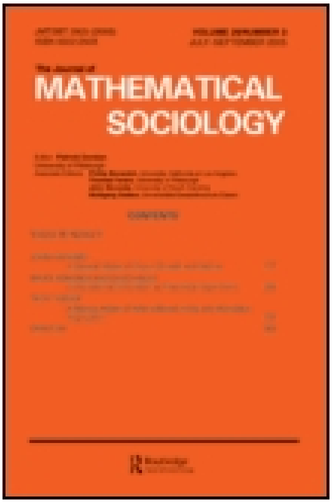

The Journal of Mathematical Sociology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gmas20

Computational Simulation as Theoretical Experiment

BRUCE EDMONDS [a] & DAVID HALES [b]
[a] Center for Policy Modelling, , Manchester Metropolitan University , Manchester, UK
[b] Department of Computer Science , University of Bologna , Bologna, Italy
Published online: 31 Jan 2007.

Journal of Mathematical Sociology, 29:209–232, 2005

ISSN: 0022-250X print/1545-5874 online
DOI: 10.1080/00222500590921283

Computational Simulation as Theoretical Experiment

Bruce Edmonds
Centre for Policy Modelling, Manchester Metropolitan University, Manchester, UK

David Hales
Department of Computer Science, University of Bologna, Bologna, Italy

Agent-based simulation can help establish the possibility and characteristics of emergent processes. However the simulation is meaningless without an accompanying interpretation. We argue that the original context needs to be carried with the simulation so as to limit excess generalization from such models. The simulation becomes a theoretical experiment which mediates between observations of the phenomena and natural language descriptions. Replication and exploration of simulations can start to identify the extent of their validity, and thus pave the way for cautions and limited generalization of results.

This is illustrated by reimplementing and re-examining two established models. Schelling's model of racial segregation is shown to give counter-intuitive results when pushed out of its intended context—the domain of valid interpretation is narrower than that covered by the whole the model. Takahashi's model of generalized exchange is shown to have included unnecessary assumptions. In this case the domain of valid interpretation is wider than the model (at least in this aspect). A tag-based variation is described where generalized exchange is shown to emerge without information about the past behavior of others.

Keywords: simulation, experiment, exchange, context, segregation, agent-based

INTRODUCTION

Formal models have long been used to inform sociological thought. In the past many of these models have often used the language of mathematics. In essence, sets of variables are related to each other via a system of equations, the values of the variables are interpreted in

Address correspondence to Bruce Edmonds, Centre for Policy Modelling, Manchester Metropolitan University, Aytoun Building, Aytoun Street, Manchester, M1 3GH, UK. Tel.: +44 161 247 6079, Fax: +44 161 247 6802. E-mail: bruce@edmonds.name

terms of the intensity or amount of a social phenomenon. In sociology, such equations have not usually been taken in a positivist sense to directly reflect real social intensities or quantities, and the fundamental difficulties of attempting this have been repeatedly pointed out, particularly in "theory of measurement" in the philosophy of science (for a survey from a social science perspective see Martin, 2003). Sometimes such systems of equations have closed-form solutions, but when this is not feasible their results are numerically simulated.

Such models have, on the whole, been used as an analogy that can be used to illuminate or illustrate ideas expressed in the richer medium of natural language. The model plays a supporting role to the main argument. In many cases, if a paper that uses such a model were rewritten without it, the conclusions would not greatly differ, but rather the formal model is a concrete expression and demonstration of the processes being discussed.

Over the past 35 years, a new kind of formal model has been increasingly used: that of the computational model. This is a model where crisp qualities are modified in a precise way by the action of specified algorithms.[1] Of course, this is not very new—in 1969 Thomas Schelling used a simple cellular automata model to demonstrate that even low levels of preference for neighbors of similar race or culture could result in emergent segregation (Schelling, 1969).[2] This model was extended into a board game with black and white counters in (Schelling, 1971).

This second type of model has been feasible as a result of the development of the computer. The computer allows such computational models to be easily built and animated in a fluid and interactive manner. Thus it allows for models to be explored and played with in a way that previously would have only been possible for skilled mathematicians.

The computer has also made it feasible to construct and use much more complicated simulations. This has made possible a type of simulation where individuals are separately represented by parts of the model. Each of these parts has its own states and can have its own algorithm; the states of these parts can be used to stand for the states of social actors and the messages passed between the parts to represent the interaction between actors. This contrasts with simulations where the actors are represented in aggregate, such as in statistical models. If the computation specific to these parts can be interpreted in cognitive terms, these parts can be called agents—thus a style of modelling called "agent-based social simulation" (Conte and Moss, 1999).

[1]Even continuous quantities are represented in a computer by a finite and crisp approximation.

[2]Schelling did not use a computer but, in essence, it was a computational simulation.

Agent-based social simulation can been seen as a move to a more transparent style of modelling (Edmonds, 2003). The relation between the model and what it stands for does not require mathematical averaging techniques (as in statistics) but can be very direct: one agent in the model stands for one actor; one message passed from one agent to another in the model stands for one communication or action; one change in the state of the agent stands for one change in the state of the actor; and so on. This, more direct, style of representation makes it far more readily comprehensible by non-experts.

Agent-based social simulation has developed from a variety of approaches (Troitzsch, 1997). This is not the place for a thorough review of its complex history; rather we direct readers to such accounts as the introduction to (Gilbert and Troitzsch, 1999) as well as (Halpin, 1999; Macy and Willer, 2002; Sawyer, 2003) and may gain a good idea of the field from the pioneering collections (Gilbert and Doran, 1994) and (Gilbert and Conte, 1995).

The transparency and interactive nature of computational simulations endow them with a persuasiveness that cold analytic models lack. This has both advantages and disadvantages. On one hand, computational simulations are more amenable to criticism in both detailed and general terms by stakeholders as well as other academics, but on the other hand they can be seductively misleading. In particular, it is often unclear with simulation models, how wide their scope is. That is, it is impossible to tell from a small set of simulation results whether these results are particular or are more widely applicable, *even when the simulation algorithm is completely specified*.

We argue that a simulation should be seen as a formal model, but one that needs to be treated, not as an analytic model, but more like a partly understood phenomena having, at best, only intermediate generality. The simulation does encode a sort of theory which, along with an interpretation, can be used to represent some aspects of social phenomena. However, this kind of theory is one that is only accessible via simulation experiments. Thus to use a simulation as a model of observed phenomena one also needs another theory—a theory of how the simulation works. Thus simulations as models of observed phenomena can fall short in at least two different ways: the theory that it is supposed to encapsulate can turn out to be inadequate, and the theory of the way in which it does this encapsulation can be mistaken.

One consequence of the fact that simulations (as used in sociology) are necessarily experimental objects is that, like other experiments, they need to be independently replicated before they can start to be trusted (Axtell et al., 1996, Edmonds and Hales, 2003).

Only in this way can one start to be confident that the simulation mechanism has been adequately understood—when a model has been independently replicated, using different computer languages, and investigated in different ways, and the results remain consistent with the theory of how it works then one can start to trust that theory (within the ranges it has been investigated). Otherwise there is always the possibility that we are mistaken about our simulations, in which case whether or not the results are consistent with what is observed does not tell us about the validity of our intended theory (which was supposed to be encapsulated by the simulation but was not). Of course, if there is a mismatch between the intended theory and how the simulation works, one might either try to fix the simulation or change the intended theory to match that of the simulation (the latter appropriate when one is using a simulation to explore ideas). In many cases the ideas that a simulation is intended to represent is left implicit (Edmonds, 2001), which makes it difficult to tell where any error lies.

In this paper, we replicate two simulations and use these replications to start to map out how far their scope extends while retaining consistency with their original interpretation. The two simulations are Schelling's model of racial segregation (Schelling, 1969, 1971) and Takahashi's model of generalized exchange (Takahashi, 2000). The point of introducing the former is that it is an example of a model whose intended scope is narrower than that of the whole model, which is demonstrated by showing that the model displays behavior counter to the original interpretation in some aspects. The latter model is introduced as an example of the opposite situation. The Takahashi model is shown to have included unnecessary assumptions in its construction, which can be dropped while retaining the results important to its intended interpretation. Thus it has a scope that is potentially wider than that of the original model.

Finally, we present a new model of generalized exchange that is a modification of Takahashi's incorporating recently discovered "tag" mechanisms (Hales, 2000; Riolo et al., 2001). We map the scope of this model and compare it with Takahashi's model. These two models are combined to throw new light onto ongoing debates.

We are making a synthetic methodological point—to highlight the role and scope of the interpretation (originating from the context of the phenomena that motivated the model) in the use of simulation models which makes explicit the relationship (particularly the differences) between the intended theory and that embodied by a simulation. We do not claim to be the first to do this—we are drawing out what has often been left implicit.

TABLE 1 Important Parameters of Schelling's Model and their Typical Values

Parameter	Values
Width and height of board	20
Proportion of grid occupied by counters, the *crowding*	{0.1, . . . ,0.9}
The critical proportion below which counters move, *c*	{0.1, . . . ,1}

EXAMPLE 1: SCHELLING'S MODEL OF RACIAL SEGREGATION

This model of Schelling is composed of black and white counters that are moved around on a 2D grid. At the start, a number of black and white counters are placed randomly about the grid. It is important that some squares are left empty. In each iteration each counter is considered in turn. For each counter, if the proportion of the eight immediate neighboring counters that are the same color as them is below a fixed critical level then, if this is possible, it moves to the nearest currently empty space where this proportion is at or above this level (measured by the number of block moves). The simulation carries on for a fixed number of iterations or until there is no more movement. We only consider cases where there are equal numbers of black and white counters and where all counters have the same level of intolerance. The main parameters are shown in Table 1.

Figure 1 shows the progress of a run where $c = 0.5$ and crowding is 0.67, after one iteration and at iteration 4 (after which no more movement occurs).

An important lesson drawn from this simulation is that, due to effects at the edges of areas of like-colored groups, these clusters develop even when the critical proportion, c, is below 50%. This was taken as indicating a possibility, namely, that racial segregation could arise as the result of a relatively low level of intolerance. Figure 2

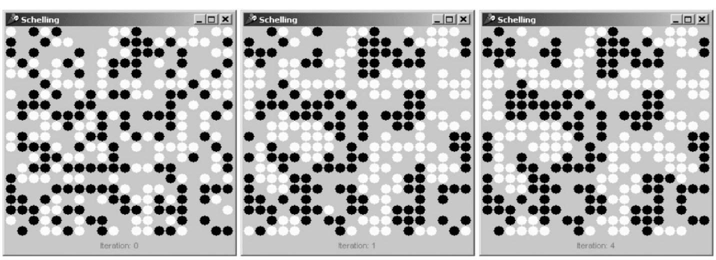

FIGURE 1 Segregation developing in the Schelling model (at iterations 0, 1 and 4).

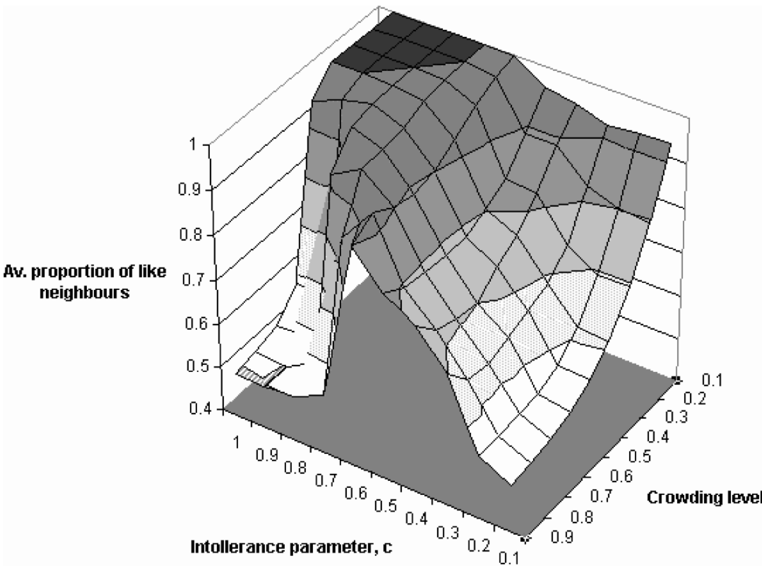

FIGURE 2 Graph of the segregation that results from different levels of intolerance and different levels of crowding in Schelling model (averaged over 25 independent runs).

shows a graph of the average proportion of like-colored counters round each counter for values of c (intolerance parameter) taken from $\{0.1,\ 0.2, \ldots, 1\}$ and levels of crowding taken from $\{0.1,\ 0.2, \ldots, 0.9\}$.

Here we see that the Schelling model results in apparently counter intuitive outcomes. At higher levels of crowding (>50%), the average level of segregation drops sharply to 50% as the level of intolerance increases towards its maximum (the bottom lefthand corner of Figure 2). This is because at high levels of crowding and intolerance there are frequently no empty spaces where the ratio of same-colored to total neighbors would be above the critical parameter, so counters do not move even though their ratio is also below the critical level. Schelling was not completely clear about what should occur if there were no acceptable spaces to move to. It seems that he only really considered cases where movement could occur; this is not surprising as is was the movement and the patterns that emerged that were his concern and the reason for discussing this model. This is *not* a flaw in Schelling's model, which still effectively shows how segregation might quickly emerge in the presence of only a weak level of intolerance. However, it does show that not *all* aspects of the original model are valid in terms of usefully informing our thinking. The model is not a *general* theory, of it was not intended to be. In other words, the model cannot be considered on its own but needs the intended interpretation to make sense.

Of course, one can change Schelling's stated rule of movement to fix this apparent problem. That is elaborate so that if there is no space where the ratio of like-colored counters are at an acceptable level, then a counter moves to the space with the highest such ratio, i.e., it seeks for any space where the ratio is like-colored counters is closer to the desired level. There is no reason to suppose that the range where the model is valid (in explanatory terms w.r.t. the intended interpretation) would happen to be exactly the possible range of parameters or settings in a simulation. This would be to conflate the intended theory with the simulation that embodies it, what we are warning against.

We present this simple case to show the utility of exploring parameters of a model (relaxing assumptions) in order to map the domain of interpretation. In this case, the model is simple enough (and well known enough) that it would be unlikely that anyone would seriously misunderstand the counter-intuitive results we pointed out (attempting to apply them to inform ideas concerning human affairs). But with more complex agent-based models, the domain of interpretation is not always so clear; therefore, we argue that this kind of analysis of existing models is a useful exercise in attempting to map their domain of interpretation. The point is *not* whether the Schelling model can be fixed in this regard—the point is that its relevant scope, is not the whole domain of possible parameter values.

In the following section, we apply this investigative mode to a more complex model concerning the emergence of generalized exchange in a small population. This turns out to be an example of the opposite sort, one where some of the assumptions built into the model turn out to be unnecessary.

EXAMPLE 2: TAKAHASHI'S MODEL OF GENERALIZED EXCHANGE

Takahashi (2000) presents two versions of an agent-based model of the emergence of generalized exchange based on agents practicing a "fairness criteria" for giving. In this section we concentrate on the non-spatial model.[3] We re-implement the model and reproduce his

[3]Takahashi's spatial model (Takahashi, 2000) is less interesting since it presents results that have been well explored and reproduced elsewhere (Nowak and May, 1992; Axelrod, 1980, 1998). Without going into details, we note that a 2D spatial layout of agents can resolve free-riding issues even *without* the need for fairness criteria-based giving.

results. We then explore a number of parameters (assumptions) to test the scope of the model. We find that the conclusions drawn from the original model are indeed robust over a number of model variants. We also identify some interesting results.

Since the focus of our paper is the experimental exploration of agent-based models (i.e., methodological) rather than a theoretically motivated sociological research question (i.e., theoretical), we give only a brief overview of the nature of generalized exchange and why it is an important issue within sociology. We refer readers with little background in the area to Takahashi's (2000) original article that gives an excellent overview of the sociological and anthropological literature and its relationship to recent computational models.

Generalized exchange forms a central topic within the classical social exchange literature (Lévi-Strauss, 1949; Malinowski, 1922). It would appear to be a major requirement in the formation and maintenance of large-scale and complex social organizations. However, theoretically, there are still major questions surrounding it: Why does generalized exchange emerge, and how is it maintained? Generalized exchange typically requires unilateral resource giving since reciprocation is not necessarily direct but from a third party. From a rational action or evolutionary perspective, unilateral giving with no guarantee of reciprocation raises the spectre of the free-rider problem—members of the exchange system may take without giving. Various possible solutions have been advanced (see Takahashi for a detailed summary). However, Takahashi notes that these previously given solutions require fixed network structures or other implausible assumptions. He advances a novel "fairness based" model of generalized exchange that he claims is more plausible and generally applicable than previous models.

The Model

The model comprises a set of agents who interact over some number of "generations." In each generation there are a number of "trial." In each trial each agent is awarded some amount of resources (NR) and then may give some amount (including none or all) of this to another agent. Each agent makes a decision as to the amount to give and to which agent to give to. In order to decide how much to give, an agent refers to its own "giving gene" (GG)[4] whose value determines the amount to give.

[4]As noted by Takahashi the biological terminology is not to be taken literally in the interpretation of the model, rather, it is used for clarity in the explanation of its mechanics.

To decide which agent to give to, each agent refers to a "tolerance gene" (TG) along with its giving gene (GG) and calculates a "giving criteria" (M) in the following manner: $M = GG \times TG$. The GG gene represents a kind of "generosity" characteristic, and the TG gene represents a "sense of fairness" level. An agent with high TG (>1) would prefer to give to those who gave *more* than it gives whereas a low TG (<1) indicates an agent that is prepared to give to others who gave less than it gives itself.

Since each agent may have different values for both GG and TG, each agent may exercise different criteria for giving. The decision of who to give to (if at all) in each trial involves each agent selecting a single agent randomly from all agents in the population that gave *at least* M resources in the last trial. If no agents meet the criteria then the agent that gave the most in the last trial is selected.[5] Then a gift of resources is made from the giving agent to the selected recipient. The resources held by the giving agent are reduced by GG, and the resources held by the recipient are increased by $GG \times RV$ (where RV is the factor by which the resources are multiplied). So when $RV > 1$, the recipient derives greater benefit from the resources than the giver would derive.[6] There are some number of trials within each generation. In each trial, each agent is awarded NR resources, calculates its M criteria, and gives as described above. When all trials have been completed, "natural selection" and "mutation" determine the members of the next generation.[7]

This involves reproducing agents based on their score (the cumulative number of resources obtained)—agents whose score was less than the group average (less the standard deviation) are called; agents with a score higher than the average (plus the standard deviation) produce two offspring and the rest produce one offspring. The GG and TG values of each offspring are subject to random mutation with probability MR (independently with probability MR the GG and TG are replaced with new random values). Table 2 shows the parameters.

Takahashi's Results

Takahashi gave results for two variants of the model. In the first (simulation 1-1) all "genes" were initialized at random, and in the second

[5]If several gave the same highest amount then one of these is selected at random.

[6]As noted by Takahashi (2000) repeating Takagi (1996), this condition has been indicated as necessary condition for the emergence of generalised exchange.

[7]The population is adjusted to ensure it is the same size over each generation.

TABLE 2 Important Parameters for the Takahashi Model

Parameter	Values
Group size	20, 100
Number of replications	50
Number of generations	1000
Number of trials per generation	10
Number of resources given to each agent, each trial (NR)	10
Giving gene (GG)	[0..10]
Tolerance gene (TG)	[0.1..2.0]
Mutation rate (MR)	0.5 (5%)
Value of each resource given to the recipient (RV)	2

TABLE 3 Comparison of Takahashi's Results with our Replication

	Takahashi's results		Our results	
	GG	TG	GG	TG
Sim 1-1	9.30 (1.25)	1.00 (0.22)	9.34 (1.21)	0.96 (0.26)
Sim 1-2	9.47 (0.65)	1.00 (0.27)	9.46 (0.79)	1.00 (0.25)

The values given are averages over 50 independent runs. The numbers in brackets are standard deviations.

(simulation 1-2) the giving gene (GG) was initialized to zero for all agents.[8] Results were calculated for 50 independent replications (with different initial pseudo-random number seeds). Each run was to 1000 generations. We reproduced the results by re-implementing the simulation model from Takahashi's description. The original and replication results are given in Table 3. These were calculated from averages of the GG and TG genes from all agents at the end of the final generation.

As can be seen, our results quantitatively match those of Takahashi.[9] This result gives us confidence that Takahashi's results are correct and that his model description is sufficiently clear. Additionally, we can be fairly confident that our re-implemented model accurately captures Takahashi's model. This latter point is important for the following experimentation with the model, where we relax various assumptions to test the robustness of the model.

[8]To test if generalized exchange would emerge when, initially, all agents give nothing.

[9]The results are not literally identical. This is to be expected due to differing implementation details (e.g., different numeric precision, pseudo-random number generators, etc.). However, the results are well within the tolerances expected at this level of replication. More importantly, the replication gives us confidence that Takahashi's results are independent of such specifics.

Testing the Robustness of the Model

To test the robustness of the model against various assumptions we conducted a number of runs, varying several parameters. First, we implemented two different reproduction methods (both consistent with Takahashi's interpretation). Takahashi's original selection and reproduction mechanism (as described above) involved culling those agents that attained a low score, producing one offspring for average scoring agents and two offspring for high scoring agents. The main interpretation here was that "those scoring well tend to be copied by those scoring less well." To test the robustness of the results with differing methods of reproduction, we implemented two additional methods (roulette wheel and random tournament) that have been widely used by others (Axelrod, 1998; Holland, 1992; Davis, 1991).

The roulette wheel method reproduces agents probabilistically based on their score. That is, in each new generation, each agent can expect to produce a number of offspring equal to its proportionate score against the whole population. So, an agent scoring three times more than another agent can expect to have three times more offspring in the next generation. Since the method is probabilistic, an element of chance (or noise) is introduced into the process.[10] The random tournament method involves selecting two agents from the population at random and reproducing the agent with the highest score to the next generation (this is repeated up to the number of required offspring for the next generation, in this case 20). We note that this latter method introduces more noise into the process, but we believe it has a more direct interpretation as a kind of social learning process.

In addition to varying the reproduction method, we also relaxed a major assumption of the original model, that is, that agents with GG values above zero *must* make a gift to another agent, even when none exists in the population that satisfies the agent's M criteria. In the original model, when an agent cannot find another who gave at least $M = GG \times TG$ in the last trial, the agent gives to the agent that gave the highest in the last trial. In order to test if this kind of "forced giving" was necessary for the emergence of generalized exchange in the model, we explored a variant in which agents do

[10]To understand this one can visualize a "roulette wheel" in which each agent occupies a number of spots on the wheel. The number of spots is proportional to the score of the agent. Reproduction involves "spinning" the wheel for each new offspring required; an agent with score P has half the chance of generating an offspring than one with a score of 2P (Davis, 1991).

not give if they can not find another agent that satisfied their M criteria.[11]

Finally, we varied the RV value, the factor by which the value of a given resource increases to the recipient. Takahashi notes that it is "standard practice" in social dilemmas research to have this value set to 2 and that it is a necessary condition that it should be greater than 1 for social exchange to emerge. We tested this assumption by varying RV from 0 to 2 in increments of 0.2 to find the point at which generalized exchange emerged.

In summary, we re-ran simulations systematically for all values of the parameters we discussed above: 3 reproduction mechanisms, 2 giving criteria (forced giving and no giving when criteria M is not satisfied) and 11 different RV values (0 to 2 in 0.2 increments). This gives a total 66 different variations. Additionally, we ran these for 20 and 100 agent populations and for random initialization of GG values and zero initialization of GG values (giving a final total of 264 variations). We made 50 independent runs (with different pseudo-random number seeds) for each variation of the parameter values giving a total of 1.32×10^4 independent runs from which we draw our analysis.

THE NEW RESULTS

The results obtained from the 264 variations have been summarized in Figure 3. This chart shows results for various parameter settings when the GG gene value is randomly initialized for each agent at the start of each run. The results produced when GG was initialized to zero are almost identical to those shown in Figure 3 and so are not reproduced here. Each line on the chart represents a different combination of reproduction method and forced or non-forced giving. The key on the chart indicates for each line the reproduction method and if giving was forced or not. As before, results are calculated from averages of the GG gene over the entire population at the end of the final generation. The results shown are for a population of 20 agents; the results obtained for a 100 agent population were almost identical to those shown so have not been included.

Overall the results follow a highly robust pattern giving us a high degree of confidence that Takahashi's observations of his model hold when various assumptions are relaxed and changed. Consider dividing the chart in Figure 3 into four quarters by splitting on the x-axis at

[11]One can imagine that this could have a dramatic effect on the results since an agent that does not give (even if it has a high GG value) will be marked as a low giver in the subsequent trial and hence potentially not be identified by others as one worthy of a gift.

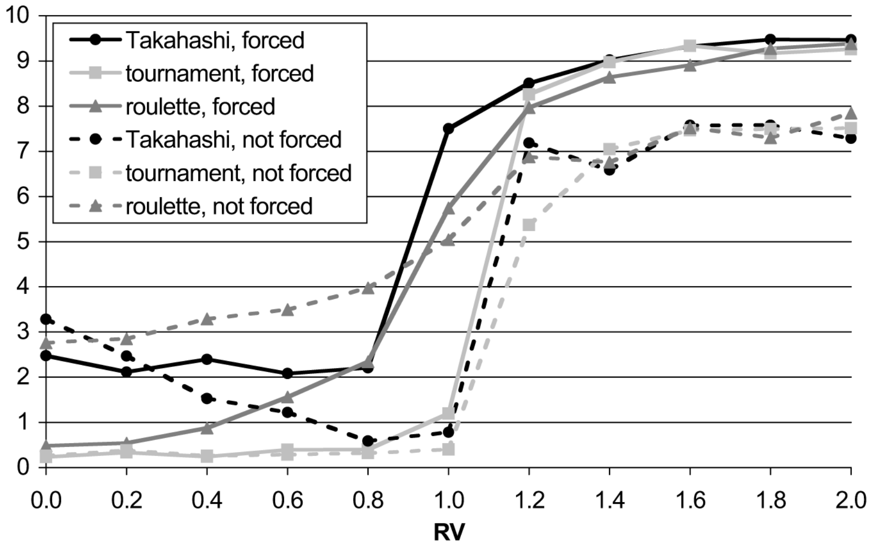

FIGURE 3 Average value of GG (at then end of the final generation) for different values of RV over different reproduction methods with forced and non-forced giving for the Takahashi model. All values are averages over 50 independent runs.

value RV = 1.0 and along the y-axis at value GG = 0.5. We note that the top-left quarter (where RV < 1.0 and GG > 5) and the bottom-right quarter (where RV > 1.0 and GG < 5) are empty, indicating that across all the variations explored, when RV > 1.0 we see the emergence of generalized exchange. This means that the conclusions drawn from Takahashi's original model are applicable to this broader set of models.

Additionally we note some further details. First, that within the top-right quarter of the chart the forced giving (solid lines) results are all higher than the non-forced giving results (dotted lines). This is not surprising since we expect that when agents are not forced to give, some agents with high GG gene values may not receive benefit from others. Most of the results in this quarter show that forced giving increases the average GG value by about 2. A more surprising result is observed when RV = 1.0. Notice that three lines are above GG = 5 on the y-axis. Specifically, the line representing Takahashi's original reproduction method combined with forced giving gives the highest result (GG > 7). What this tells us is that if Takahashi's original model is changed so that RV = 1.0 (i.e., that the receiving agent gets the same value as that which was given), then we still find the emergence of some generalized exchange.

You may recall our previous hypothesis concerning such a possibility. Takahashi repeated a claim by Takagi (1996) that a necessary condition of social exchange was that $RV > 1.0$. However, using Takahashi's original assumptions we have found that this is not strictly the case. We therefore conclude that either there is something wrong with the claim or something important missing from this original version of the model in relation to Takagi's assertion. However, as can be seen from the results, only half (3) of the model variants produce $GG > 5$ when $RV = 1.0$. This indicates that only under specific kinds of learning (reproduction methods) and giving criteria (*mainly* non-forced in these results) does social exchange emerge when resources have no additional value for recipients.[12]

In the bottom-left quarter of the chart we found some unexpected results.[13] In this quarter we see what happens when a gift loses value (i.e., the recipient receives less than was given). We see a wide variation of GG value but broadly the GG value either increases as RV increases or stays much the same. However, for non-forced giving using Takahashi's reproduction method as RV increases the GG value decreases significantly (until $RV = 1.0$). This is due to the ineffectiveness of Takahashi's reproduction method when there is almost no variation in the population, e.g., when most have exactly the same score except a couple who do worse.

Discussion of New Results

When we initially embarked on the process of relaxing and changing the assumptions of the model we expected to "break" the model. That is, from our experience with other similar models, we expected the results to be highly sensitive to these assumptions. However, we were surprised to find that the majority of Takahashi's results generally hold for our model variants, although we identified some interesting quantitative differences. With this in mind we argue that *these results strengthen and generalize Takahashi's original findings.* Fairness based giving under perfect information of past giving can emerge generalized exchange in a wide range of model variants. We claim that this kind of replication and extension of models is one of the major contributions that agent-based modelling has to make to sociological theorizing.

[12]We feel this would be a very interesting line of future work, developing a theory relating various factors to the emergence of exchange without any increase factor.

[13]Indeed, so puzzling to us initially that we independently re-implemented the model in another computer language, and the results produced were identical.

EXAMPLE 3: GENERALIZED EXCHANGE WITHOUT KNOWLEDGE OF PAST

Behavior or Imposed Social Structure

As discussed above, Takahashi's model of the emergence of generalized exchange requires that agents have the ability to base their giving of resources on the past giving behavior of others. The model gives all agents access to perfect information concerning what all other agents gave in the last trial. Takahashi therefore positions his model within a set he calls "discriminating altruism" under perfect information of past behavior. However, he identifies that although his model is more general than previous models based on dyadic reciprocal altruism such as the Tit-For-Tat strategy in the Prisoner's Dilemma (Axelrod, 1980, 1998) or indirect reciprocity model (Nowak and Sigmund, 1998), it becomes less credible when applied to large-scale societies when agents interact with others about whom they have no previous information (i.e., exchange with *complete strangers*). It seems unconvincing that each agent would have access to the previous behavior of all other agents in a large-scale society. Takahashi refers to the "stranded driver" scenario, in which it is not considered unreasonable for the driver of a broken-down car to elicit help from passing strangers. Indeed such help seems to be provided often in many societies where it is not considered abnormal to elicit or provide help in such situations. One way to explain such behavior is by appeal to a pre-existing norm of behavior. But (as noted by Takahashi) how can such a norm become established in the first place if we assume self-interested agents? Here we present a new model that is a modification of Takahashi's model but *requiring no information of past interactions or a pre-existing fairness norm*. The model applies recently discovered mechanisms based on "social cues" or "tags" (Holland, 1993) applied to interactions with strangers without an imposed social structure (Hales, 2000, 2001; Riolo et al., 2001).

These provide a novel spin on "tribalism" or "in-group" altruism (Hardin, 1982). Rather than fairness-based discrimination (as in Takahashi's model) our model utilises social similarity based discrimination. Agents only make gifts to agents that they judge to be sufficiently similar to themselves (what we interpret as their "in-group"). Although generalized exchange in our model only occurs within the in-group, the group boundaries are highly permeable since they are based on socially learned cues or "display traits" (tags) which are learned and mutated in exactly the same way as the giving behaviors themselves (Allison, 1992; Holland, 1993; Macy and Skvoretz, 1998). This simple mechanism can produce high amounts of giving between strangers by providing a means by which social structure can develop.

The Model

Our model is essentially the same as Takahashi's model (see above); however, we dispense with the tolerance gene TG. Instead each agent stores a giving gene (GG) and a set of binary tags. Each tag represents the presence or absence of some observable social cue.[14] Initially the GG genes and the tags for each agent are initialised randomly. In the same way as Takahashi's model, the simulation is executed for some number of generations and in each generation some number of trials are performed. In each trial each agent is awarded some amount of resource, and based on the value of its GG gene attempts to give some amount to another agents. However, in this model agents do not search the population based on a "fairness criteria" since they have no knowledge of what the other agents gave in the last trial (they are strangers); rather, the agent simply searches the population for an agent that posses exactly the same tag values as itself. If it finds such an agent, it makes a gift by deducting an amount equal to GG (the agents giving gene) from its own account and passing this to the receiving agent. The receiving agent receives the gift amount multiplied by RV (as in Takahashi's model). If no agent can be found then either no gift is made (in the non-forced gifting model variant) or a gift is made to a randomly chosen agent whatever their tag values (in the forced gifting model variant).

This assumption of exact tag matching enforces strict boundaries between agents sharing the same tags partitioning the population into distincting groups. This is not a strict requirement of this kind of cooperative process; however, previous work has experimented with "tolerance-like" distance metrics or partial matching schemes with "dont care" bits (Rolio et al., 2001; Hales, 2001). Recent work has applied a similar mechanism for agents embeded in a "social-like" network interacting with nearby others (Hales, 2004).

At the end of each generation the population is reproduced based on the score of each agent in the same way as described previously (we explore the same three reproduction variants). During reproduction both the GG gene value and the tag values are copied to offspring and mutation is applied to each (i.e., mutation is applied to the GG

[14]These might include any characteristic that can be observed and learned socially, including style of dress or manner of speech. Bowles and Gintis (2000) use this representation in their analytical model of parochialism.

TABLE 4 Major Parameters for the Tag Model of Generalized Exchange

Parameter	Value
Group size	20, 100
Number of replications	50
Number of generations	1000
Number of trials per generation	1
Number of resources given to each agent, each trial (NR)	10
Giving gene (GG)	[0..10]
Number of binary tags associated with each agent	32
Mutation rate (MR)	0.001
Value of each resource given to the recipient (RV)	[0..2]

gene and to each tag value independently with probability MR[15]). In this way, the reproduction process can be interpreted as a kind of social learning based on imitation in which both the giving gene GG and the tag values are copied from higher scoring agents to lower scoring agents.

It is important to note that the tag values have no linkage to the GG values—their only function in the model is to act as social cues. They are initialized at random and are (at least initially) completely arbitrary. Any association they come to have with particular GG values results from the evolution of the model. This can be contrasted with the Tags (called "social cues") used in Macy and Skvoretz (1998) in which some of the cues were related to the actual cooperation behavior of the agent (or "character"). No stipulation is made in this model that any particular GG value implies any particular tag values, the evolutionary process alone emerges any such associations.

The major parameters of the tag model are shown in Table 4. The larger population size (100; however, see later results for 20 agents), lower mutation rate (0.001) and large number of tags (32) are values that have been "imported" from work with a previous tag model (Hales, 2000) in which it was found (via a systematic search of the parameter space) that these values promoted the altruistic tag mechanism (this is discussed later). Note also that we only have one trail per generation. The reason for this is efficiency because in the tag model there is no knowledge of past behavior so only one trail is required—further trials will produce the same results.[16]

[15]When mutation is applied to a tag bit it is "flipped" (i.e., 0 to 1 or 1 to 0).
[16]No significant difference was found when we increased the number of trials.

The Results

Figure 4 shows the results produced from our tag model for each of the different variations of reproduction method and forced or non-forced giving (as described for Figure 3). Each point shows the average GG value of the population at the end of the final generation (averaged over 50 runs).

The chart shows a number of interesting characteristics. Most strikingly we observe that only two lines occupy the top-right quarter of the chart indicating that generalized exchange only emerges for two combinations of reproduction method and non-forced giving. Note that none of the forced giving models as emerged generalized exchange. Also notable is that non-forced giving combined with Takahashi's reproduction method produces a flat line around GG-5. When we investigated this we found that this resulted from random drift. That is, no giving occurs at all in this condition—no gifts are ever made because the reproduction mechanism is not sufficient to produce agents with identical tags (tag clones). This is due to the strict reproduction mechanism. Takahashi used, only producing multiple offspring if agents had a significantly higher score than the average. In the other reproduction mechanisms there is a higher stochastic

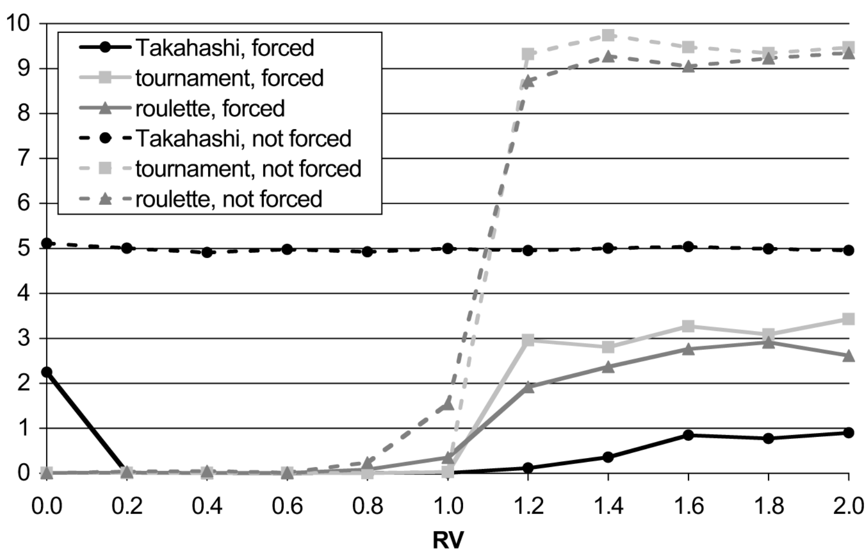

FIGURE 4 Average value of GG (at then end of the final generation) for different values of RV over different reproduction methods with forced and non-forced giving for the Tag model for 100 agents. All values are averages over 50 independent runs.

element that allows agents to produce multiple offspring even if they are not significantly higher scorers, which is a necessary condition for this model to emerge generalized exchange. This implies that our model actually benefits from noise in the reproduction.

All forced gifting model variants (whatever the reproduction mechanism) languish in the bottom-right quarter of the chart. This means that none of them has produced high amounts of generalized exchange. We conclude therefore that a further necessary condition for the emergence of generalized exchange in our model is that gifting should not be forced, that is, that agents should have the ability *not* to give if they do not find a suitable partner to give to (in this case, a partner with the same tag values). This contradicts the generality of findings from a previous model (Hales, 2000) where tags were sufficient to produce high cooperation in single-round Prisoner's Dilemma interactions where interaction was forced when partners with matching tags could not be found. We note that in this case Takagi's (1996) hypothesis holds in all the runs with our tag model. Where high levels of exchange emerge this is only when RV $>$ 1.0.

Discussion of Results

Let us be clear what our tag model is demonstrating in the cases where high levels of exchange *did* emerge: *self-interested agents are emerging high levels of exchange without knowledge of past interactions of others*. We consider this to be interesting—it goes against much of the received wisdom concerning the behavior of self-interested actors within an evolutionary setting.

In both Hales (2000) and Riolo (2001), the mechanism by which tags promote this kind of behavior has been discussed (see Hales, 2001, for an in-depth discussion). Briefly, the process operates thus:

- Via differential reproduction groups of agents form with identical tags.
- Groups with more in-group gifting increase the scores of its members.
- High-scoring groups increase in size due to score-based reproduction.
- Groups are invaded by free riders (which do not give) by mutation.
- The invaders multiply in the group due to their higher score than co-operators (givers) until the group is dominated by free-riders.
- A group full of free-riders destroys itself since they gain no benefits from giving to each other.

It is important to realize that the process does not stop freeloading; rather, freeloaders destroy the tag groups (agents sharing the same tag values) that they occupy by reducing the scores of the other agents within them. So long as some other group exists that does not contain freeloaders, agents from that group will ultimately outperform groups containing freeloaders. Of course, mutation means that at some stage freeloaders will invade any group that initially does not contain them. So the dynamics of the model follow a continual process of group formation, invasion by freeloaders and dissolution. Altruistic in-groups constantly blink in and out of existence as generations pass.

The tag model therefore never reaches any kind of "equilibrium state," where there is high and consistent donation (as does the Takahashi model). In the scenarios where the tag model produces high overall gifting a constantly dynamic group formation and dissultion process is occurring. We believe this might be more plausable when compared with many large scale human societies where dynamics are observed over time (consider for example the formation, growth and dissolution of corporations or firms over time).

We note that the low mutation rate when applied independently to each tag bit and the GG gene effectively means that an agent's tag "as a whole" has a higher chance of changing (as tag are mutated) than the GG gene. It would seem that this kind of arrangement is required for the tag process to produce altruism, since this allows time for tag groups containing agents with high GG gene values (i.e., altruists) to spread to other tag groups (by tag mutation) before the group is invaded by a freeloader (by mutation to the GG genes).

However, our investigation has shown that this emergence of exchange does not occur when gifting is forced (i.e., when agents are forced to give resources even if they cannot find a partner with the same tag values) in the current model. Also, we have found that Takahashi's reproduction mechanism applied to our tag model does not produce exchange because it does not produce multiple offspring from a single agent when agent scores are identical.

We therefore view these results as productive since they circumscribe the generality of the interpretation of these kinds of tag mechanisms. We appear to have identified two *necessary conditions* for such tag mechanisms to operate, which have not been identified in previous tag models (Hales, 2000; Hales, 2001; Riolo et al., 2001). We therefore ring a note of caution over extravagant interpretations of such models applied to human societies (Sigmund and Nowak, 2001).[17] It would

[17]The interpretations given by Sigmund and Nowak (2001) relate to Riolo et al.'s (2001) tag model. However, on-going work (Edmonds and Hales, 2003; Roberts and Sherratt, 2002) has identified that this model (even when fixed) effectively collapses to the mechanism given here.

appear that before such interpretations can be made into specific human social contexts, we would need to be clear as to the nature of the social learning occurring there and be confident that we have a reproduction mechanism that captures it.

We "imported" several of our assumptions (number of tags, mutation rates, population size) from previous work (Hales, 2000) and view the success of the tag mechanism in producing high levels of exchange as evidence that results from one agent-based model can be productively applied to another. We note that this will be an increasingly important activity if agent-based modelers are to benefit from each other's findings and progress (discussed later).

In Hales (2000) we performed a search of part of the parameter space of the model, and the results of that exploration suggest that our tag exchange model would be sensitive to a reduction in the number of tags, number of agents or number of generations, or an increase in mutation rate. The tag processes are not merely applicable to large-scale societies but *require* large-scale societies to operate. Indeed our previous results (Hales, 2000) showed that the larger the society the more quickly cooperation became established. Future work could explore the effect of varying these parameters within our new tag model.[18] An initial experiment, reducing the population size to 20 (see Figure 5), in line with Takahashi's original model, shows that cooperation is significantly reduced from the 100 agent case (see Figure 4). Interestingly, the basic profile of the results remains with just the higher cooperation selection mechanisms being reduced. This result is to be expected with such a low population size since the "group selective" process can only operate if there are enough agents in the population to form a number of competing groups.

This final results indicate that the Takahashi model would be more applicable to situations in which small groups produce stable long-term "giving norms" by each agent, knowing the past behavior of each other agent and interacting regularly. The Tag model applied to much larger "anonymized" societies in which there is a perpetual and dynamic group formation dissolution process occurring. There seems no reason to suppose that in a given target social situation *both* models could be applicable at different scales. However, without empirical investigations such speculations can only be posed as conjectures.

[18]Initial experimentation with individual runs for lower numbers of agents and numbers of tags has confirmed this. We found that with less than 80 agents and 25 tags, exchange disappeared in the current model. However, more runs needed to justify a more confident claim.

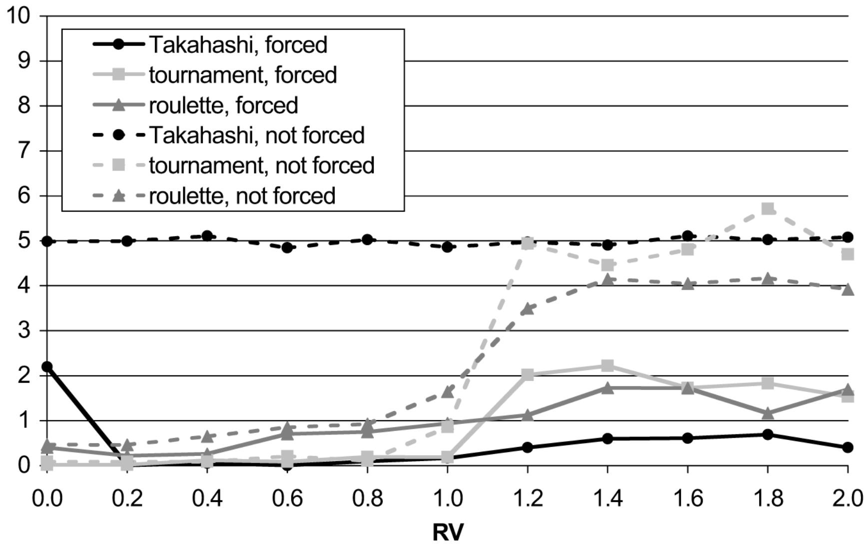

FIGURE 5 Average value of GG for different values of RV for the same parameters of the Tag model as given in Figure 4 but for a population of 20 agents.

CONCLUSION

Often in sociology one needs to be able to demonstrate the possibility and characteristics of emergent processes, in particular, global results that arise out of the interaction of many actors by way of a process that is far from obvious. A simulation can help do this, along with an interpretation of that simulation. A simulation without an interpretation is meaningless, and a simulation without a good interpretation is useless. The interpretation gives meaning to the simulation by reference to particular context in the world of social phenomena. This paper hopes to illustrate how this context needs to be brought in to simulation via the interpretation. Results from a simulation are specific to the model and are generalized to form meaningful interpretations, but there are distinct limits as to how far results can be generalized.

In fact, we would argue that good agent-based simulation is particularly appropriate for informing thought on social processes, because it only generalizes from the phenomena to a modest, appropriate extent (when used properly). The simulation is sufficiently complex to be able to mimic some of the complexity of social phenomena but formal enough to be reliably experimented upon. In a meaningful sense the simulation mediates between the abstract ideas of theories expressed in nature language descriptions and the observations of the phenomena. However, this does require some care—we are naturally adept at using and judging analogies. We can appreciate their point while being aware of the dangers inherent in their unthinking application—we need to be as vigilant with simulations. To do this we need to do (at least) two things: be aware of (and preferably make explicit) the scope of the simulation (i.e., the bounds under which it is thought to be faithful to its interpretation) and establish a norm that simulations need to be replicated and explored by independent researchers. If these practices become established, then agent-based simulation offers the promise of a sociology in which representations of theories (agent-based models) can be replicated, explored, and related to produce a more or less consensual body of knowledge.

REFERENCES

Allison, P. (1992). The cultural evolution of beneficent norms. *Social Forces*, *71*(2), 279–301.

Axelrod, R. (1980). *The Evolution of Cooperation*, New York: Basic Books.

Axelrod, R. (1998). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press.

Axtell, R., Axelrod, R., Epstein, J. M., & Cohen, M. D. (1996). Aligning simulation models: A case study and results. *Computational and Mathematical Organization Theory*, *1*, 123–141.

Bowles, S. & Gintis, H. (2000). Optimal Parochialism: The Dynamics of Trust and Exclusion in Networks. *SFI Working Paper* 00-03-017. Santa Fe Institute, Santa Fe, N.M.

Conte, R. & Moss, S. (1999). Special interest group on agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, *2*(1), http://jasss.soc.surrey.ac.uk/2/1/4.html.

Davis, L. (1991). *Handbook of Genetic Algorithms*, New York: Van Nostrand Reinhold.

Edmonds, B. (2001). The use of models—making MABS actually work. In S. Moss & P. Davidson, (eds.), *Multi Agent Based Simulation* (pp. 15–32). Springer: Lecture Notes in Artificial Intelligence, 1979.

Edmonds, B. (2003). Against: a priori theory for: Descriptively adequate computational modelling. In *The Crisis in Economics: The Post-Autistic Economics Movement: The First 600 Days* (pp. 175–179). Routledge.

Edmonds, B. & Hales, D. (2003). Replication, replication and replication – some hard lessons from model alignment. *Journal of Artificial Societies and Social Simulation*, *6*(4), http://jasss.soc.surrey.ac.uk/6/4/11.html.

Gilbert, N. & Conte, R. (1995). *Artificial Societies: The Computer Simulations of Social Life*, London: UCL Press.

Gilbert, N. & Doran, J. E. (1994). *Simulating Societies: The Computer Simulations of Social Phenomena*, London: UCL Press.

Gilbert, N. & Troitzsch, K. G. (1999). *Simulation for the Social Scientist*, Buckingham, UK: Open University Press.

Hales, D. (2000). Cooperation without space or memory: Tags, groups and the Prisoner's Dilemma. In S. Moss & P. Davidson, (eds.), *Multi Agent Based Simulation* (pp. 157–166). Springer: Lecture Notes in Artificial Intelligence, 1979.

Hales, D. (2001). *Tag Based Co-operation in Artificial Societies*, Ph.D Thesis, Department of Computer Science, University of Essex, UK.

Hales, D. (2004). From selfish nodes to cooperative networks. In *Proceedings of the Fourth IEEE International Conference on Peer-to-Peer Computing* (P2P2004), 25–27 August 2004, Zurich, Switzerland. IEEE Computer Society Press.

Halpin, B. (1999). Simulation in sociology. *American Behavioral Scientist*, *42*, 1488–1508.

Hardin, G. (1982). Discriminating Altruisms. *Zygon*, *17*, 163–86.

Holland, J. (1992). *Adaptation in Natural and Artificial Systems*, Cambridge, MA: The MIT Press.

Holland, J. (1993). The Effect of Labels (Tags) on Social Interactions. *SFI Working Paper* 93-10-064. Santa Fe Institute, Santa Fe, N.M.

Lévi-Strauss, C. (1949). *Les Structures Elementaires De La Parente*, Paris: Presses Universitaires de France.

Macy, M. W. & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, *28*, 143–166.

Macy, M. W. & Skvoretz, J. (1998). The evolution of trust and cooperation between strangers: a computational model. *American Sociological Review*, *63*, 638–660.

Malinowski, B. (1922). *Argonauts of the Western Pacific*, London: Routledge.

Martin, O. (2003). Philosophical and scientific origins of the representational theory of measurment (1930–50). *Social Science Information Sur Les Sciences Sociales*, *42*, 485–513.

Nowak, M. & May, R. (1992). Evolutionary games and spatial chaos. *Nature*, *359*, 826–929.

Nowak, N. & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.

Riolo, R. (1997). The effects of tag-mediated selection of partners in evolving populations playing the iterated Prisoner's Dilemma. *SFI Working Paper* 97-02-016, Santa Fe Institute, Santa Fe, N.M.

Riolo, R., Cohen, M. D., & Axelrod, R. (2001). Cooperation without reciprocity. *Nature*, *414*, 441–443.

Roberts, G. & Sherratt, T. (2002). Does similarity breed cooperation? *Nature*, *418*, 499–500.

Sawyer, R. K. (2003). Multiagent systems and the micro-macro link in sociological theory. *Sociological Methods and Research*, *31*, 325–363.

Schelling, T. C. (1969). Models of segregation. *American Economic Review*, *59*, 488–493.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, *1*, 143–186.

Sigmund, K. & Nowak, M. A. (2001). Tides of tolerance. *Nature*, *414*, 403–405.

Takagi, E. (1996). The generalized exchange perspective on the evolution of altruism. In W. B. G. Liebrand & D. M. Messick, (eds.), *Frontiers in Social Dilemmas Research* (pp. 311–336). Berlin: Spring-Verlag.

Takahashi, N. (2000). The emergence of generalized exchange. *American Journal of Sociology*, *105*, 1105–1134.

Troitzsch, K. G. (1997). Social simulation – origins, prospects, purposes. In R. Conte, R. Hegselmann, & P. Terna, (eds.), *Simulating Social Phenomena* (pp. 41–54). Lecture Notes in Economics and Mathematical Systems, 4565. Berlin: Springer.