

# Computational Simulation as Theoretical Experiment

*Bruce Edmonds and David Hales*

([bruce@cfpm.org](mailto:bruce@cfpm.org) and [dave@davidhales.com](mailto:dave@davidhales.com))

Centre for Policy Modelling, Aytoun Building, Manchester Metropolitan  
University, Manchester, M1 3GH, UK.

Tel. +44 (0) 161 247 6074

## 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 terms of the intensity or amount of a social phenomenon. In sociology such equations have not 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. Sometimes such systems of equations have closed-form solutions but when this is not feasible their results are simulated.

However 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 only plays a supporting role to the main argument. In many cases if a paper that uses such a model were re-written without it, the conclusions would not greatly differ. This is also why the use of analytic solutions (as opposed to numerical simulations) is not greatly important to their use.

Recently a new kind of formal model is being 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. Of course, this is not very new - back in 1969 Thomas Schelling used a simple cellular automata model to demonstrate that even low levels of preference for neighbours of similar race or culture could result in emergent segregation. This model resembled a board game with black and white counters being moved around a chequerboard according to fixed rules and throws of a die.

This second type of model has been greatly facilitated (some would say empowered) by 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 such 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 possible 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 of the program to represent the interaction between actors. If the computation of these parts can be interpreted in cognitive terms these parts are often called 'agents'. Thus we have a style of computational model called agent-based social simulation.

Agent-based social simulation can thus be seen as a move to a more transparent and accessible style of modelling. 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 in the model makes it far more readily comprehensible.

The transparency and interactive nature of such models endow them with a persuasiveness that 'cold' analytic models do not have. This has both advantages and disadvantages. On one hand they are more amenable to criticism in both detailed and general terms, 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 algorithm of the simulation is completely specified*.

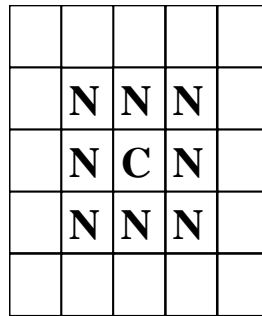
We argue that a simulation should be seen as a formal model of intermediate generality, but one which needs to be treated not as an analytic model but more like a partly understood phenomena. The simulation does encode a theory which, along with an interpretation can be used to represent some aspects of social phenomena. However the nature of this theory is one that is only accessible via the experimentation of running the simulation. Thus to use a simulation as a model of observed phenomena one also needs a theory of how the simulation works.

One consequence of the fact that simulations (as used in sociology) are necessarily experimental objects is that, like other experiments, they need to be replicated before they can start to be trusted. In this paper we replicate two simulations, and use these replications to start to map out how far their scope extends whilst retaining consistency with their original interpretation. The two simulations are Schelling's model of racial segregation (Schelling 1969) and Takahashi's model of generalised exchange (Takahashi 2000). Finally we present a new model of generalised exchange that is a modification of Takahashi's incorporating recently discovered novel "tag" mechanisms (Hales 2000, Riolo 2001). We also map the scope of this model and compare it with Takahashi's original model. In this latter section we demonstrate how results from different models can be combined into hybrids that throw new light onto on-going debates.

## **Example 1: Schelling's model of racial segregation**

This model is composed of black and white 'counters' which are moved around on a large chequerboard. At the start an number of black and white counters are placed randomly about the board. It is important that a good number of squares are left empty (say a third of them). The counters are permanently but randomly numbered, so that 1 is allocated to a randomly chosen counter, 2 to another etc. In each iteration each counter is considered in turn in order of their numbering. For each counter, if the proportion of their immediate neighbouring counters that are the same colour as them is below a fixed critical level then if there are any empty spaces in its neighbourhood then it chooses

one such space randomly and moves to it. The neighbourhood of a counter is all those locations 1 square away including diagonals (Figure 1). The simulation carries on as long as is desired (e.g. until there is no more movement).



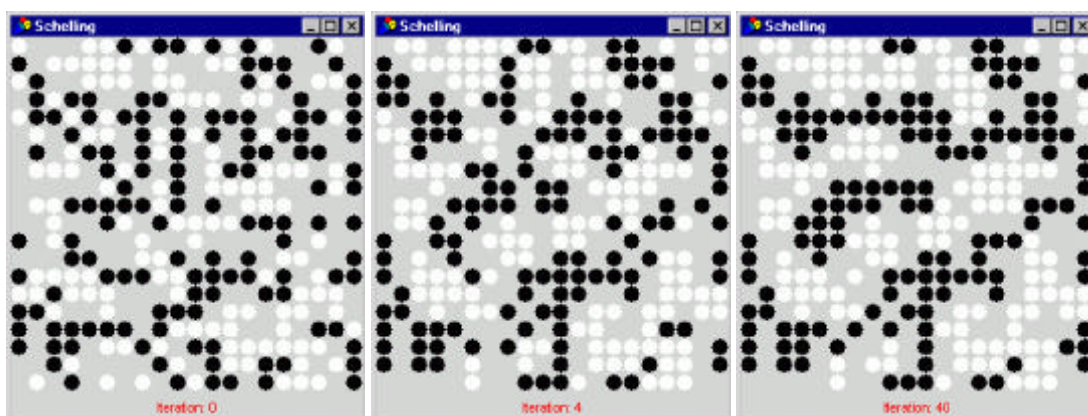
**Figure 1. The neighbourhood of a counter in the original Schelling model (a distance of 1 from the counter)**

The principle parameters of the model (along with my default parameters) are shown in Table 1 below.

Parameter	Default value
Width of board	20
Height of board	20
Number of black counters	133
Number of white counters	133
The critical proportion below which counters move	0.5
Distance that defines neighbourhood	1

**Table 1. Important parameters of Schelling's model**

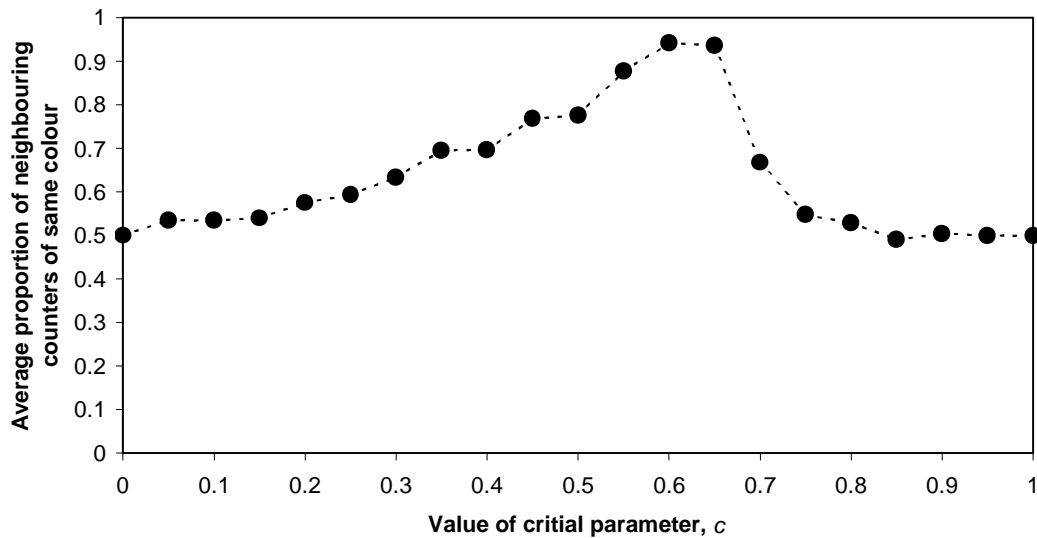
Figure 2 below shows the progress of an example simulation for the default parameter values above.



**Figure 2. Segregation emerging in the Schelling model (at the start, iteration 4 and iteration 40 respectively)**

The important lesson drawn from this simulation is that due to effects at the edges of areas of like coloured groups, these clusters develop even when the

critical proportion,  $c$ , is below a half. This was taken as indicating a possibility: namely that racial segregation could arise as the result of a relatively low level of intolerance.



**Figure 3. Graph of the segregation that results from different levels of intolerance with a 1-distance neighbourhood**

However, the Schelling model demonstrates an initially counter intuitive result which is evidenced in Figure 3. Here a number of runs were been executed, for different values of the critical parameter ( $c$ ). The y-axis shows a measure of segregation in the final board arrangement measured as the average proportion of neighbours of the same colour (averaged over all counters). Notice that initially as  $c$  is increased (i.e. as each agent becomes more intolerant of different coloured neighbours) the amount of segregation increases. However, when  $c > 0.65$  the amount of segregation in the final arrangement on the board decreases back to around 0.5.

These results seem initially to be counterintuitive<sup>1</sup>. Reflection on the nature of the model reveals what's happening. Since movement is to a *randomly* chosen free square within the neighbourhood of a counter, very high levels of  $c$  produce so much movement that no clusters of segregation have a chance to form.

This is clearly not representative of what would happen if there were such a high level of mutual intolerance in the real world - people might seek to escape areas with even low levels of people they were intolerant of, but they would not relocate randomly, they would seek to move to an area that was "better" in their opinion. This is not a serious flaw in Schilling's model, which still effectively demonstrates that segregation might emerge in the presence of only a weak level of intolerance, due to effects at the edges of segregated areas. However it does show that it is only an aspect of the total model that is valid in terms of

<sup>1</sup> Further runs in which the size of the neighbourhood was increase to a distance of 2 and 3 squares on the board produced broadly similar results.

usefully informing our thinking - the model is not a general theory. In other words the model can not be considered on its own but needs the context of the intended interpretation to make sense.

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 have presented (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, we therefore argue that this kind of analysis of existing models is a useful exercise in attempting to map their domain of interpretation. In the following section we apply this investigative mode to a more complex model that claims to capture the emergence of generalised exchange in a small population.

## **Example 2: Takahashi's model of generalised exchange**

Takahashi (2000) presents two versions of an agent-based model of the emergence of generalised exchange based on agents practicing a "fairness criteria" for giving. In this section we concentrate on the non-spatial model<sup>2</sup>. We re-implement the model and reproduce his results. We then explore a number of parameters (assumptions) to test the robustness 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 and counter-intuitive results from this exploration.

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 very brief overview of the nature of generalised 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 it's relationship to recent computational models.

Generalised 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 organisations. However, theoretically, there are still major questions surrounding it: Why does generalised exchange emerge, and how is it maintained? Generalised 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 dreaded spectre of the free rider problem – members of the exchange system may take without giving. Many and various possible solutions have been advanced (see Takahashi for a detailed summary). However, Takahashi notes that these previously given solutions require fixed

---

<sup>2</sup> 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-rider issues even *without* the need for fairness criteria based giving.

network structures or other implausible assumptions. He advances a novel “fairness based” model of generalised 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 “trials” in which each agent is awarded some amount of resources (NR) and given the opportunity to give some amount (including none or all) of these resources to other agents. Each agent makes a decision as to the amount to give and to which agent to give to (selecting a single receiving agent in each trial). In order to decide how much resource to give, an agent refers to a personally stored “giving gene” (GG)<sup>3</sup>. The value of the giving gene indicates the amount to give. Each agent may have a different giving gene.

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 when giving. The decision of who to give to (if at all) in each trial involves each agent calculating its criteria value M and 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<sup>4</sup>. When a recipient agent has been selected then a gift of resources is made from the giving agent to the 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 some factor by which the resources are multiplied). So when  $RV > 1$ , the recipient derives greater benefit from the resources than the giver would derive<sup>5</sup>. Within each generation there are some number of trials. For each trial each agent is selected, awarded NR resources, calculates its M criteria and executes the process of gifting as described above. When all trials have been completed, “natural selection” and “mutation” determine the members of the next generation<sup>6</sup>.

---

<sup>3</sup> 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 the mechanics of the model.

<sup>4</sup> If more than one agent jointly gave the same highest amount in the last trial then one is selected at random.

<sup>5</sup> As noted by Takahashi (2000) repeating Takagi (1996), this condition has been indicated as necessary condition for the emergence of generalised exchange.

<sup>6</sup> For simplicity the population is adjusted to ensure it is the same size over each generation.

Parameter	Default value
Group size	20
Number of replications	50
Number of generations	1000
Number of trials per generation	10
Number of resources given to each agent in 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 2. Important parameters for the Takahashi model.**

This involves reproducing agents based on their score (i.e. the cumulative number of resources obtained – either as gifts or directly). Specifically, agents whose score was less than the group average (less the standard deviation) are culled; agents with a score higher than the average (plus the standard deviation) produce two offspring and agents with a score between these values produce a single offspring. The GG and TG values of each offspring are subject to random mutation with probability MR<sup>7</sup>. Table 2 gives the values of the important parameters for the Takahashi model.

## Takahashi's Results

Takahashi gave results for two variants of the model. In the first (simulation 1-1) all “genes” were initialised at random and in the second (simulation 1-2) the giving gene (GG) was initialised to zero for all agents<sup>8</sup>. 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 computer simulation model from Takahashi's description of it. The original and replication results are given in Table 3. The results were calculated from averages of the GG and TG genes from all agents at the end of the final generation.

---

<sup>7</sup> More precisely: independently with probability MR the GG and TG are replaced with new randomly generated values from the acceptable range.

<sup>8</sup> The idea here was test if generalised exchange would emerge from a group in which, initially, all agents gave nothing.

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

**Table 3. Comparison of Takahashi's results with our replication. The values given are averages over 50 independent runs. The numbers in brackets are standard deviations.**

As can be seen, our results quantitatively match those of Takahashi<sup>9</sup>. This result gives us confidence that Takahashi's results are correct and that his model description is sufficiently clear. Additionally, we can be 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.

### Testing the robustness of the model

In order to test the robustness of the model against various assumptions we conducted a number of runs, varying several parameters. Firstly, 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". In order 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<sup>10</sup> (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<sup>11</sup>. The random tournament method involves selecting two agents from

---

<sup>9</sup> The results are not literally identical. This is to be expected due to differing implementation details (e.g. different numeric precision, different pseudo-random number generators, the variance of independent runs etc.). However, the results are close enough to be well within the tolerances expected at this level of replication. Importantly, the replication gives us confidence that Takahashi's original results are independent of such implementation specifics.

<sup>10</sup> Takahashi's method is rather specific and to our knowledge it has not been used elsewhere in evolutionary models. This rang alarm bells, so our natural reaction when testing robustness was to apply more generally used reproduction methods.

<sup>11</sup> To understand the process one should visualise 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 to the next generation involves "spinning" the wheel for each new offspring required (in this case 20 times for a new population of 20 agents). This means that an



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 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 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 generalised exchange in the model we explored a variant in which agents do not give if they can not find another agent that satisfied their M criteria<sup>12</sup>.

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 generalised exchange emerged. We hypothesize that if generalised exchange did not first occur for a value of RV at 1.2 then either the model misses some important mechanism or the assumption is incorrect.

In summary, we re-ran simulations systematically for all values of the parameters we have 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 of  $3 \times 2 \times 11 = 66$  different variations. Additionally, we ran these for random initialisation of GG values and zero initialisation of GG values (giving a final total of 132 variations)<sup>13</sup>. We made 50 independent runs (with different pseudo-random number seeds) for each variation of the parameter values.

## The New Results

The results obtained from the 132 variations have been summarised in Figure 4. This chart shows results for various parameter settings when the GG gene value is randomly initialised for each agent at the start of each run. We note that the results produced when GG was initialised to zero was almost identical to those shown in Figure 4 and so are not reproduced here. Each line on the chart represents a different combination of reproduction method and forced or

---

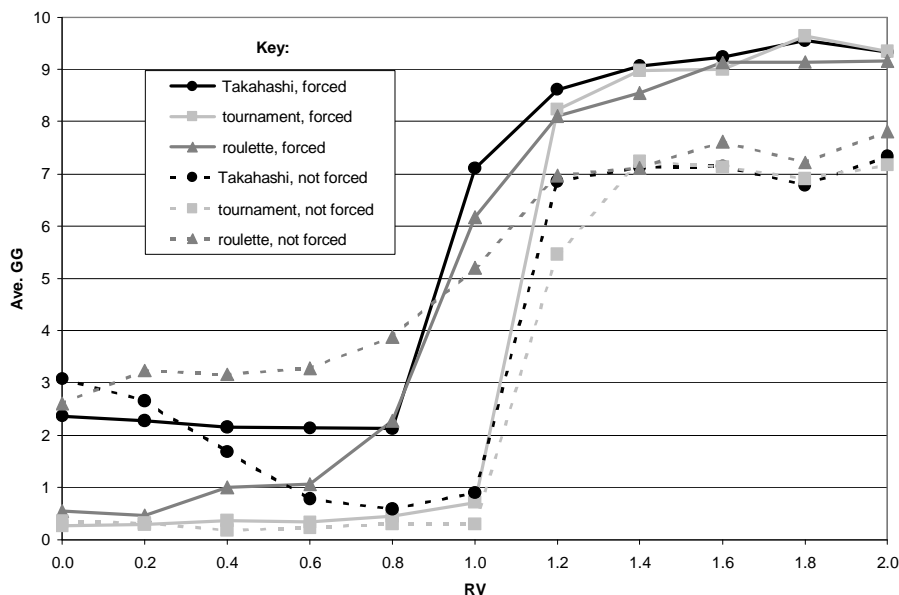
agent with score P has half the chance of generating an offspring on each spin than an agent with a score of 2P (see Davis 1991 for more details).

<sup>12</sup> 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.

<sup>13</sup> Since each variation was executed 50 times (with different pseudo-random number seeds) the results presented are a synthesis of 6,600 individual runs. At the micro-level each run is for 20 agents to 1000 generations with 10 trials per generation so the results are the product of 1,320,000,000 individual agent decisions.

non-forced giving. The key on the chart indicates for each line the reproduction method and if giving was forced or not. The three reproduction methods are Takahashi's original method, the random tournament method and the roulette wheel method (each described previously). As before, results are calculated from averages of the GG gene over the entire population at the end of the final generation.

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 4 into four quarters by splitting on the x-axis at 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 (more or less) of generalised exchange. This is good news for Takahashi's original model, since it means that the conclusions previously draw from it are applicable to this broader set of models.



**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 Takahashi model. The key shows the reproduction method and if forced or non-forced giving was used. Takahashi's original model used forced giving. All values are averages over 50 independent runs.**

Additionally we note some further details. Firstly, that within the top-right quarter of the chart (where  $RV>1.0$  and  $GG>0.5$ ) the forced giving (solid lines) results are all higher than the non-forced giving results (dotted lines). This is a non-surprising result since as speculated earlier 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 of resources as that which was given) then we still find the emergence of generalised exchange (albeit at a slightly lower level than when  $RV = 2.0$ ).

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 gain no additional value to the recipient of a gift<sup>14</sup>.

In the bottom-left quarter of the chart we found some quite puzzling and unexpected results<sup>15</sup>. 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  $GG$  values 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  $RV=1.0$ ). We currently are unsure of the mechanism behind this effect<sup>16</sup>.

### 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 generalise Takahashi's original findings*. Fairness based giving under perfect information of past giving can emerge generalised exchange. This claim holds over 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.

## Example 3: Generalised exchange *without* knowledge of past behaviour

---

<sup>14</sup> 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.

<sup>15</sup> Indeed, so puzzling to us initially, that we independently re-implemented the model in another computer language – the results produced were identical.

<sup>16</sup> Although of only passing interest to the conclusion of this paper we again feel that a greater understanding of this result would be an interesting line of future work. Our general meta-theoretical metric being: *if it's unexpected it's worth investigating*.

As discussed in the previous section, Takahashi's model of the emergence of generalised exchange requires that agents have the ability to base their giving of resources on the past giving behaviour 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 behaviour. 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 models (Nowak and Sigmund 1998) it becomes less credible when applied to large-scale societies when agents interact with others that they have no previous information about (i.e. exchange with *complete strangers*). It seems unconvincing that each agent would have access to the previous behaviour 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 behaviour is by appeal to a pre-existing norm of behaviour. 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" applied to interactions with strangers (Hales 2000, 2001, Riolo 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 (their in-group). Although generalised 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 behaviours themselves (Allison 1992, Holland 1993). This very simple mechanism can produce high amounts of giving between strangers.

## 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<sup>17</sup>. 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

---

<sup>17</sup> These might include style of dress, manner of speech or any observable characteristic that can be observed and learned socially. Bowles and Gintis (2000) use this same tag representation in their analytical model of parochialism. We prefer to use binary characteristics since they appear to offer a more intuitive interpretation (i.e. the presence or absence of some observable and socially learnable trait).

awarded some amount of resource and based the on the value of it's 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. 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).

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 gene and to each tag value independently with probabality MR<sup>18</sup>). 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.

<b>Parameter</b>	<b>Default value</b>
Group size	100
Number of replications	50
Number of generations	1000
Number of trials per generation	1
Number of resources given to each agent in 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]

**Table 4. Major parameters for the tag model of generalised exchange.**

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 initialised at random and are (at least initially) completely arbitrary. Any association they come to have with particular GG values is an emergent property from the evolution of the model.

The major parameters of the tag model are shown in Table 4. The larger population size (100), low mutation rate (0.001) and large number of tags (32)

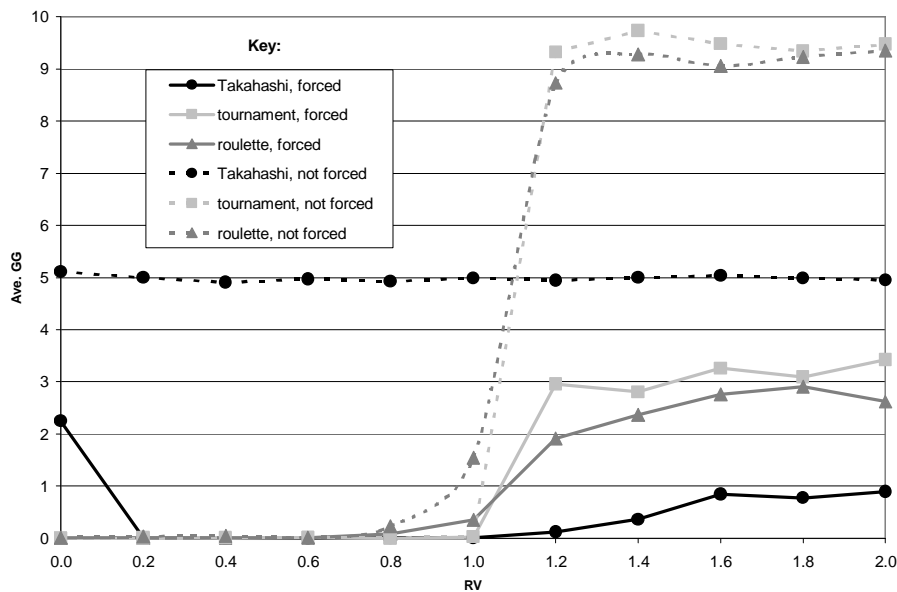
---

<sup>18</sup> When mutation is applied to a tag bit it is "flipped" (i.e. replacing a zero with a one or a one with a zero). As previously, when mutation is applied to the GG gene, a new GG value is selected at random from the range [0..10].

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 simply efficiency – since in the tag model there is no knowledge of past behaviour only one trial is required.

## The Results

The chart shown in Figure 5 details 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 4 in the previous section). As before each point on the chart shows the average GG value of the population at the end of the final generation (averaged over 50 independent runs).



**Figure 5. 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. The key shows the reproduction method and if forced or non-forced giving was used. All values are averages over 50 independent 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 generalised exchange only emerges for two combinations of reproduction method and non-forced giving. Note that none of the forced giving models as emerged generalised 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 element that allows for agents to produce multiple offspring

even if they are not significantly higher scorers – which is a necessary condition for our tag model to emerge generalised exchange<sup>19</sup>.

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 generalised exchange. We conclude therefore that a further necessary condition for the emergence of generalised exchange in our model is that gifting should not be forced - 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<sup>20</sup>.

We note that in this case Takagi's (1996) hypothesis holds in all the runs with our tag model. Where high levels of exchange do 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 (i.e. with strangers).*

We consider this to be a rather astounding result that overturns much of the received wisdom concerning the behaviour of self-interested actors within an evolutionary setting.

In both Hales (2000) and Riolo (2001) the mechanism by which tags promote this kind of altruistic behaviour between self-interested agents 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 tag values
- Groups with more in-group gifting increase the scores of in-group members
- Groups with high scores increase in size due to reproduction based on score
- Groups invaded by free riders, which do not give, reduce the scores of the other members of the group and effectively destroy the group

It is important to realise that the process does not stop freeloading; rather freeloaders destroy the tag groups (agents sharing the same tag values) that

---

<sup>19</sup> Interestingly, this implies that our tag model actually benefits from noise in the reproduction mechanism.

<sup>20</sup> However, in that model, runs were to 100,000 generations and cooperation did not emerge until several thousand generations, future work could explore longer runs. We are not optimistic that this would change our results since in Hales (2000) cooperation dominated immediately from the first cooperative act – in this model cooperative (giving) acts are occurring but not increasing in the population to produce high-levels of GG.

they occupy by reducing the scores of the other agents within them. So long as some other group exists which does not contain freeloaders, agents from that group will ultimately out-perform 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 constant process of group formation, invasion by freeloaders and dissolution. This process constantly repeats. Altruistic in-groups constantly blink in and out of existence as generations pass.

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 (one or more tag bits being mutated) than the GG gene. It would seem that this kind of arrangement is a requirement to drive the tag process to produce altruism, since this allows for tag groups containing agents with high GG gene values (i.e. altruists) to have time to spread the altruistic behaviour to other tag groups (by mutation of the tags) 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 can not 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 scores are more-or-less identical across all agents in the population.

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 2001). We therefore ring a note of caution over extravagant interpretations of such models applied to human societies (Sigmund and Nowak 2001)<sup>21</sup>. It would 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 in those contexts 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 modellers are to benefit from each other's findings and progress (we discuss this later in the conclusion).

---

<sup>21</sup> We should note that the interpretations given by Sigmund and Nowak (2001) relate to Riolo et al's (2001) tag model. However, on-going work (Edmonds and Hales forthcoming, also see Roberts and Sherratt 2002) has identified that this model (even when it is modified to work correctly) effectively collapses to the same clone donation mechanism given here. We note that Riolo et al use a tournament reproduction method but appear to generalise the interpretation to any kind of social learning.



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<sup>22</sup>. Future work could explore the effect of varying these parameters within our new tag model<sup>23</sup>.

## Discussion

Despite the fact that Schelling's, Takahashi's and our tag model were relatively simple, they all had the capacity to give counter-intuitive results when pushed beyond the limits of their intended context of interpretation.

In Schelling's model, segregation of counter colours failed to emerge given a high enough level of intolerance, due to the large amount of random movement that resulted. This model could have easily been fixed by assuming that when a counter is dissatisfied with the composition of its neighbours, it would randomly choose a location but only move if the new location were preferable to it. This would have ensured that random movement did not disrupt any developing segregation and hence ensure that high level of intolerance resulted in high degrees of segregation. However, this 'fix' would have ruined the point of the model, which was (by our reading) to show how segregation can arise under very weak pressures of intolerance: that of moving randomly only a little if one only had a few similar neighbours. That segregation could arise under stronger pressures of intolerance is obvious - there is no need of a model to demonstrate this.

Takahashi's model included an unnecessary assumption (that of forcing gift even when the agent's giving criteria were not met), which it did not need in order to demonstrate the authors point. Replicating the model revealed that this can be safely removed, so that the model is strengthened, whilst still making its point. Additionally, by applying a number of reproduction mechanisms the model was further strengthened in it's domain of interpretation (it would seem that the general conclusions draw from the model are robust to various models of social learning).

In our tag model we identified two candidate necessary conditions<sup>24</sup> for the emergence of generalised exchange in a scenario very similar to Takahashi's. However, we demonstrate that generalised exchange can emerge *without* information about the past behaviour of others in the populations. We consider

---

<sup>22</sup> Future work could test this hypothesis for our tag exchange model presented here.

<sup>23</sup> Some initial experimentation with individual runs for lower numbers of agents and numbers of tags confirmed this. We found that with less than 80 agents and less than 25 tags exchange disappeared in the current model. However, more runs need to be performed to make a more confident claim.

<sup>24</sup> We identify them as "candidate" conditions because without a fuller understanding of the model and further exploration we can not at this stage make a stronger claim.

this to be an important contribution to the debate on the emergence of generalised exchange in large-scale societies.

Also, our results allow us to circumscribe interpretations that have been made based on previous tag models that, we believe, are revealed to be over generalised from models that have been only cursorily explored and understood. We see this as a major danger when working with agent-based models but hope that the work presented in this paper, demonstrates (at least partially) a way to avoid such pitfalls.

## 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 a 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 generalised to form meaningful interpretations, but there are distinct limits as to how far results can be generalised.

In fact, we would argue that good agent-based simulation is particularly appropriate for informing thought on social processes, because it only generalises 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 natural 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 whilst 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 (that is, the bounds under which it is 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 new kind of sociology in which theoretical constructs (agent-based models) can be replicated, explored, related and cross-fertilised to produce a more-or-less consensual body of theoretical 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.
- Bowles, S. and 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.
- Davis L. 1991. *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold.
- Hales, David. 2000. "Cooperation without Space or Memory: Tags, Groups and the Prisoner's Dilemma." In *Multi-Agent Based Simulation*, edited by Moss, S and Davidsson, P. LNAI 1979. Berlin: Springer-Verlag.
- Hales, David. 2001. *Tag Based Co-operation in Artificial Societies*. PhD Thesis, Department of Computer Science, University of Essex, UK.
- Hardin, Garret. 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, NM.
- Lévi-Strauss, C. 1949. *Les structures elementaires de la parente*. Paris: Presses Universitaires de France.
- Malinowski, B. 1922. *Argonauts of the Western Pacific*. London: Routledge.
- Nowak, M. and May, R. 1992. "Evolutionary Games and Spatial Chaos." *Nature* 359:826-929.
- Nowak, N. and 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, NM.
- Riolo, R., Cohen, M. D. & Axelrod, R. 2001. "Cooperation without Reciprocity." *Nature* 414, 441-443.
- Roberts, Gilbert and Sherratt, Thomas. 2002. "Does similarity breed cooperation?" *Nature* 418, 499-500.
- Schelling, T. 1969. "Models of Segregation." *American Economic Review* 59:488-493.
- Sigmund and Nowak. 2001. Tides of tolerance. *Nature* 414, 403-405.
- Takagi, E. 1996. "The Generalized Exchange Perspective on the Evolution of Altruism." Pp. 311-36 in *Frontiers in Social Dilemmas Research*, edited by W. B. G. Liebrand and D. M. Messick. Berlin: Spring-Verlag.
- Takahashi, Nobuyuki. 2000. "The Emergence of Generalized Exchange". *American Journal of Sociology* 105:1105-1134.