11 Shaping social environments with simple recognition heuristics

Peter M. Todd and Annerieke Heuvelink

Imagine walking through the dark streets of Berlin on a cold night, looking for a place to get a good hot chocolate. You've been to this neighborhood only a couple of times before, so while you have experienced a few of the bars, you don't know much about them. You can't see in through the steamy windows, so you just have to make a choice and go in, hoping it won't be one of those places where the music screeches to a halt and all the locals look up from their hot chocolates to glare at you as you step inside. How can you decide which to try? You could risk a choice at random, or choose one of the places you already recognize; or you could call up a friend or two and ask for recommendations. But if your phone is broken and you can't communicate with anyone who has more knowledge, you could also think back to your previous experience and recall how many other people were in the bars at that time, or even how many acquaintances had been there when you were there. What would happen to the popularity of the bars if you and everyone else used one of these methods to choose where to drink chocolate? Would all bars be equally visited, or would some become very popular while others floundered? Will your decisions create hotspots and dead zones, shaping the social environment of Berlin bars?

There is good reason to believe that your choices will indeed shape the fate of chocolate-purveying scene, rather than just maintaining the status quo. The combined decisions of a population of agents can powerfully shape their environment, often leading some things–people, novels, cities, etc.–to become much more well-known, and more widely preferred or chosen, than others. This can be seen in the J-shaped function relating the popularity or success of items in some domain to their rank in that domain–for instance, best-selling authors sell vastly more books than the great majority of little-known authors, and the most popular consumer brands, from soft drinks to soap, sell much more than their lesser competitors (Hertwig, Hoffrage, & Martignon, 1999).

But how does such agreement, in the form of a great number of people making the same choices, come about? What psychological mechanisms underlie this cultural structure?

Much research in anthropology and human behavioral ecology has gone into showing that some simple psychological mechanisms can evolve and help people to find and converge on beneficial cultural innovations in a variety of settings (see Henrich & McElreath, 2003, for an overview, and Boyd & Richerson, 1985, for details). In particular, prestige-based mechanisms direct people to copy the behaviors and choices of successful individuals, while conformity-based mechanisms specify determining the most common behaviors and choices in a population and following those. Both types of mechanisms can for example lead all the hunters in a group to adopt bows rather than blow-darts, or the farmers in a region to plant potatoes rather than corn.

However, even simpler cognitive mechanisms, which neither seek to identify successful models nor keep track of frequencies of behaviors in a population, can enable a population of interacting individuals to coalesce strongly on a few cultural options, as seen in modern environments despite the vast number of choices available. As we will demonstrate in this chapter, just making choices based on the options one recognizes can lead to population convergence of the same sort seen in the anthropological models with two options of differing quality—even when there are many options, and when they all have the same underlying quality (as in the many bars in Berlin all getting their hot chocolate from the same underground pipe). This convergence relies on the recognition knowledge of individuals arising through their interactions with others, either through communication or indirect observation. Thus we argue that simple recognition-based decision mechanisms operating in a social setting may achieve some of the same culture-shaping effects as the "biased transmission" mechanisms explored previously.

1 Making decisions using recognition

While we humans may pride ourselves on our ability to make intelligent choices in a challenging world, people are limited in the amount of information we can process, the amount of time we can process it in, and the amount of computation our minds are able to carry out. For most of our decisions, we rely on simple cognitive heuristics, shortcuts that enable us to make good-enough choices quickly and cheaply. The surprising finding of a growing body of psychological research is that such "fast and frugal" heuristics can exploit the structure of information in the task environment to make decisions that are as good as, and in some cases better than, what more complex and information-hungry mechanisms would produce (Payne, Bettman, & Johnson, 1993; Gigerenzer, Todd, & the ABC Research

Group, 1999). These simple components of our mind's "adaptive toolbox" (Gigerenzer, 2001), some of which are evolved and some of which are learned, are thus by virtue of their fit to the environment often the best tool available for a particular inferential job.

Perhaps the simplest decision heuristic is the recognition heuristic (Goldstein & Gigerenzer, 1999, 2002), which actually makes use of an individual's *lack* of knowledge. It is based on the deep-rooted cognitive capacity to remember and recognize (rather than recall) particular names, faces, locations, and objects. The recognition heuristic can be used by agents who do not know anything about a set of options they must choose between, other than whether they have encountered each particular option before or not. The heuristic then simply says to select one of those options that are recognized (in a binary fashion, yes or no) over those that are not. If there is more than one available option that is recognized, then the recognition heuristic chooses randomly among them; if none of the options are recognized, then a completely random choice is made. Thus, the recognition heuristic can only be used when the decision-maker knows about some of the objects in a particular set, but is ignorant of others.

People and other animals use the recognition heuristic in a variety of settings. For instance, Norway rats use recognition knowledge gained by smelling the breath of their nestmates to guide their food choice on subsequent foraging trips, preferring to sample recognized foods (Galef, 1996). In laboratory settings, people use the recognition heuristic to decide which of two cities is larger, or which of two rivers is longer, or which of two sports teams wins more often (Goldstein & Gigerenzer, 2002). Furthermore, these recognition-based decisions are highly accurate when the heuristic is used in a domain where recognition and ignorance are appropriately structured, that is, where objects higher on the criterion (e.g. length of river) are more often recognized. This is likely to be the case whenever objects that are extreme on some criterion dimension are more often talked about among individuals or mentioned in the media. Goldstein and Gigerenzer (1999, 2002) showed how this holds for the city-size dimension: Large cities are more often mentioned in newspaper headlines than small ones, which in turn can drive the greater recognition for larger cities that makes the recognition heuristic *ecologically rational* to use in this task environment.

Recognition knowledge is often a highly valid cue to the structure of the environment, giving the recognition heuristic high rates of inferential accuracy. And people are sensitive to this power of recognition: In small-group settings, when the goal is to agree upon a particular decision such as which of two cities is larger, those individuals who can use the recognition

heuristic—that is, those who recognize one city but not the other—are often given more influence than others who know both cities (Reimer & Katsikopoulos, 2004).

When all of the options to be decided among are recognized by an individual, then the recognition heuristic as just presented cannot be used to choose between them. However, there can still be informative differences in the recognition knowledge for each option—some things may have been encountered more recently, or more often, than others, and so may have a higher overall activation in memory. Something like this memory activation probably underlies the recognition judgment in the first place: As long as it is above a particular threshold, the object is judged "recognized", and if the activation is below the threshold, the object is "unrecognized" (Schooler & Hertwig, in press). The recognition heuristic throws away any differences in activation values that are above the "recognized" threshold, but another heuristic, Schooler and Hertwig's fluency heuristic, capitalizes on those differences to choose the highest-activated option. This strategy works well at selecting options that have been more commonly encountered in the environment, and is thus ecologically rational when objects that are higher on the choice criterion are also more often experienced. Schooler and Hertwig have furthermore shown that both the recognition and fluency heuristics benefit from a particular amount of forgetting, so that recognition memory does not become clogged with every object ever encountered, no matter how far back in time. (See also Todd & Kirby, 2001, for the importance of forgetting in agent-based recognition models of the sort investigated here.)

Where do we get the knowledge stored in recognition memory? It can come from individual experience, encountering different objects or behavioral options as we move about in the world, and storing the more commonly encountered things more strongly in memory through more frequent and recent updates. But our recognition knowledge can also come from others. We can directly hear about things that our conspecifics recognize, as when a friend tells us about a great new bar she's just been to, and again the more we hear about some thing, the stronger is its activation trace in memory. The strength of recognition memory could in addition be influenced indirectly through social interaction, without communication of knowledge between individuals. Particular options or behaviors may be activated more highly if we see that others have made the same choice. Much work has been done on social conformity to explore how information about the decisions of others can sway one's own decisions. Asch (1956) showed that people would change their judgments of line-length when others around them made obviously-wrong judgments. This social conformity

increased as the number of others increased up to 5; bigger groups did not increase the conformity effect much further. Another study done by Milgram, Bickman and Berkowitz (1969) also showed this effect of group size. Milgram had confederates look up into the sky in the streets of New York City, with the greater the number of confederates, the more people passing by who would also stop and look up, ranging for 4% with one confederate up to 84% of passers-by with 15 confederates.

Latané (1981) summarized such results in the Law of Social Impact in order to explain the effect of groups of people on single individuals. His Law states that the total impact of a group on a single target person will increase with the strength of the group members, their number, and their proximity to the focal individual in time and space. Strength can be authority, but it can also be familiarity–you are more likely to conform to people you are close to socially than to strangers. Such factors underlie some of the model-based cultural learning mechanisms discussed by Henrich and McElreath (2003), such as focusing on social models similar to oneself. Here we will concentrate on the effect of the number of other individuals who have made the same choice as oneself. (Similar effects of the influence of other conspecifics on an individual's behavior can be seen in other species–see Noble & Todd, 2002, for connections.)

The spreading of ideas (whether in the form of knowledge, memes, fads, products, etc.) through societies has also been studied from other perspectives over the years, ranging from conditional decision models (in which the decisions of individuals are based on the decisions made by others—see Granovetter, 1978) in sociology to the use of statistical mechanics for modeling socioeconomic interactions (Durlauf, 1997). Economists have developed models to explore what determines the eventual share of a product in a certain market. Arthur (1988) proposed that when self-reinforcing spreading mechanisms are present in an economic system, common features will arise. These features include the existence of multiple equilibria in terms of what ideas or products will ultimately be adopted (different asymptotic market-share "solutions" are possible, so that the outcome is not uniquely predictable), possible inefficiency (if one idea is inherently "better" than others, but has "bad luck" in gaining early adherents, the eventual outcome may not be of maximum possible benefit), lock-in (once an equilibrium is reached, it is difficult to exit from), and path-dependence (the early history of market shares—in part the consequence of small events and chance circumstances—can determine which idea prevails). More recently, the increased interest in networks and their structure has led to new research on the spreading of ideas or

products from a more sociological perspective, taking into account the structure of the social networks that individuals find themselves in. This work addresses questions such as in which network structures ideas spread fastest, or which nodes should be targeted in order to get an idea adopted (Grönlund and Holme, in press).

In the models we present here, some of these features can arise, but others are currently not present because of our simplifying assumptions; for instance, because we assume equal fitness of the spreading items, the phenomenon of inefficiency cannot arise. In our minimalist approach we also do not incorporate preexisting social networks (though networks can be observed as emergent aspects of the agents' interactions in our models). Furthermore, we do not include more complex features such as varying expectations or personalities of the modeled individuals, as we aim to show that even much simpler processes can give rise to strong spreading patterns. (For a more complex model that incorporates such aspects, see Lane, 1997.)

2 Methods—Agent-based models for simulating social decisions

To investigate how decision-making agents can shape their environment in a coordinated fashion without direct communication, we built a family of agent-based simulation models in NetLogo. In these models, agents inhabit a world full of locations that they can choose to visit, and each agent maintains a memory of locations they have seen, as well as in some cases of other agents they have seen. As agents build up knowledge about their world and use it to decide where to go, we watch for whether their decisions combine to create new structure–hotspots and dead zones in how agents are spread across locations–in their environment. Note that in these simulations, we assign all agents to use a particular decision mechanism and see how that affects the structure of the environment they help create, rather than looking at the evolution and spread of a particular decision or learning mechanism through the population as done by other modelers (e.g., Boyd & Richerson, 1985; Henrich & McElreath, 2003).

We look for the emergence of environment structure in these simulations in two main ways: The distribution of how many patches or locations are chosen by different numbers of agents can vary from a Poisson distribution in which most patches are chosen by only one or two agents—the unstructured environment in which our models start, shown in Figure 1—to a situation where a few patches are currently chosen by many agents (e.g. 9 or 10), and are

known (recognized) by nearly all of the agents—a clumpy world where knowledge and choices are focused on a small subset of the possibilities, shown for example in Figure 2b. We also track the correlation between how often they patches visited, or chosen, by agents and how well they are known—in other words, the correlation between choice and recognition, or behavior and knowledge. If there is indeed co-evolution of the knowledge about the environment in terms of who knows what, and the structure of the environment in terms of who decides to go where, we expect this correlation to rise.

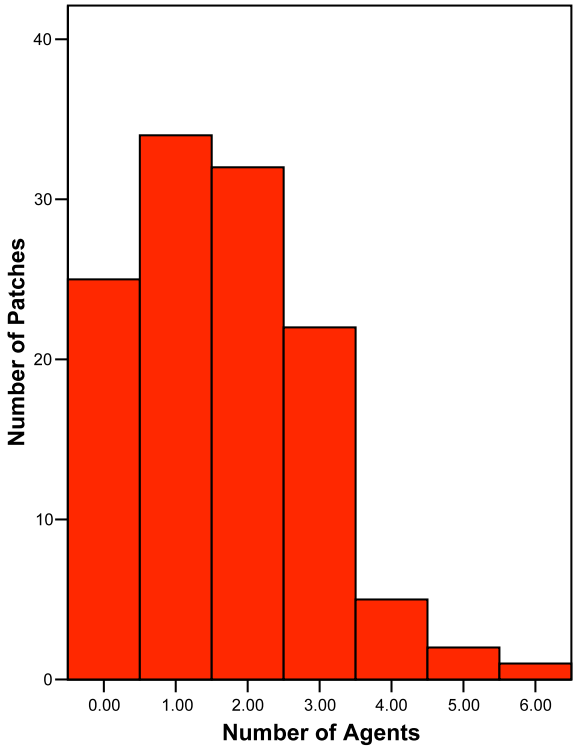

Figure 1a. Histogram of the number of patches (y-axis) that are known by a certain number of agents (x-axis) showing the near-Poisson distribution of agents randomly scattered in the unstructured environment at time step 1. (Most patches contain a single agent.)

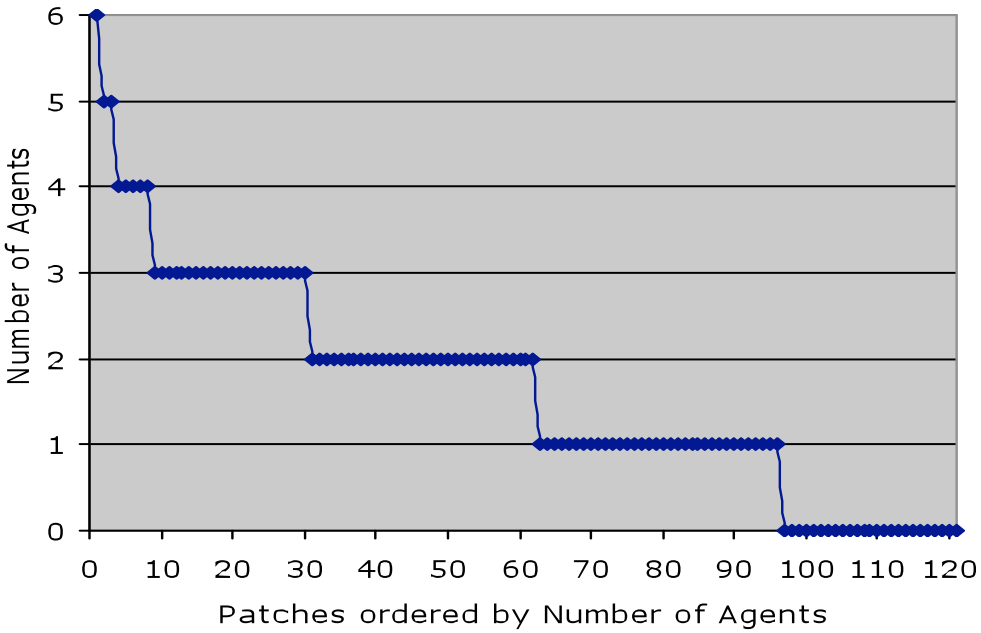

Figure 1b. Another view of the distribution of agents (now on y-axis) among the patches (on x-axis) at time step 1, with patches rank-ordered from left to right by number of agents present. Note the absence of a strong J-shaped distribution, indicating an unstructured (clumpless) social environment.

For the sake of speed we started off exploring the different models with 121 patches forming the 11x11 environment and with 200 agents forming the population. Incorporated in all our models are a memory for places and a memory for other agents for every agent. The program starts by randomly scattering the agents into the environment. At the beginning all the patches and agents are homogeneous. This is important, because we want to explore how one place can be known more and consequently visited more than another without there being any underlying difference between them (e.g., the difference in bar attendance should not be explainable by some secret ingredient that one puts in its hot chocolate).

As the simulation runs, at every time step each agent is presented with a choice between four patches they can go to. Each agent makes a decision among these somehow, using a rule or heuristic applied to their current knowledge. As mentioned in the introduction, humans can make decisions simply by looking at how well they recognize the options. We use this same recognition heuristic to let our agents decide to which of the presented options they want to go. They recognize a place if it is in their memory. They can pay attention to that

information in two different ways. The first is strictly binary: Do I recognize this option, yes or no? This binary recognition knowledge gets used by the recognition heuristic described earlier: Agents always go to a place (i.e., choose an option) they recognize. If they recognize more than one place, a decision is made at random between the recognized options. When none of the options is recognized the agent selects one at random. The second way to use recognition knowledge is as a continuous variable: How well do I recognize this option? This real-valued knowledge is used by the fluency heuristic: Agents always go to the place they recognize best. If no options are recognized or multiple options are recognized with the same value (which is unlikely), the decision is made at random between the tying options.

After deciding where they want to go every agent moves to its selected patch and increases the activation (if any) of that patch in its memory with a value of 1.0. (Note that this and most of the other parameter values are arbitrary, with the rough differences between them being more important than their precise values. The goal is to see whether any reasonable settings of the parameters will lead to the emergence of social environment structure.) As indicated earlier, forgetting is also an important component of this memory model; here, the memory trace of each patch simply decays (falls) by a default value of 0.1 with every time step. If the memory trace for a certain patch falls below zero, that patch is no longer remembered. (Thus the recognition threshold for this rudimentary memory model is 0.0; any positive memory trace results in the object being judged as "recognized".)

A second important aspect in our models is the attention the agents pay to other agents around them. As indicated earlier, people are readily influenced by others, more so the more familiar they are with those others. In our family of models we explore different ways in which agents pay attention to other agents. All the agents have a memory for other agents they meet, updated on each encounter with a default value of 1.0, just as in the memory for locations. This memory trace also decays every time step with a default value of 0.1.

We begin with a default model in which agents do not pay any attention to other agents, only to the patches they visit. Next, we look briefly at the effect of allowing individuals to communicate with each other about the patches they recognize. Finally, we consider two models with indirect social influence in which agents pay attention to the other agents they encounter, in the following ways: First, individuals can pay attention to how many other agents are on the current patch. In this case this patch is stored in their memory with the default value plus a certain value for every other agent on that patch. Think about a person walking into a bar and finding a lot of people inside. That person will deduce that this

is a quite popular bar and remember it as a good place to go. Second, individuals can notice the agents they recognize in the current patch, and can use this agent knowledge to modify how strongly they store their experience of the current location. In particular, they remember the patch they are on with the default value plus a certain increment for every other agent on that patch that they recognize. Imagine again the person walking into a bar and seeing a few others there that she recognizes from other popular bars she goes to–that's an indication that the current bar is also a happening spot to frequent.

3 Results—When does environment structure emerge?

To look for the emergence of environment structure with various direct and indirect forms of social influence and sharing of knowledge between agents, we ran a number of models according to the variations just described, with agents using either the recognition heuristic with binary memory values (models marked "bin") or the fluency heuristic with continuous memory values (models marked "con"). All the results shown here are the average results for ten runs after 20,000 time steps (to allow each model to reach a more or less steady state). For detailed presentation and discussion of these results, see Heuvelink (2004).

3.1 Model 1 — Agents on their own

The two models in which the agents pay no attention to the other agents in their environment do not produce emergent environment structure; instead, the distribution of agents over locations (chosen options) remains much the same as at the beginning of the simulation, still distributed randomly and creating a Poisson distribution. The correlation between the number of agents at a location and the number of agents that recognize that location is about 0.3. This is in part because agents often do not have the opportunity to use their choice mechanism. On average agents know about 10 out of 121 patches, because they store each patch they visit with a value of 1.0 and this value decays by 0.1 every time step. Thus, in $(1 - 9.9 / 121)^4 \approx$ 71% of their choices, agents recognize none of the four options and have to choose between them randomly. Furthermore, all the patches end up being known by similar numbers of agents on average.

We can change this by giving the agents a longer memory (lower decay) and letting them remember more locations. But even when they recognize 50 locations, and so can use the recognition heuristic about 90% of the time, still no environment structure emerges. Why not? After all, the agents must be more likely to end up at certain patches–the ones they

know–compared to other patches. The problem here is that all the agents have their own set of options that they recognize. This set is personal, local to each individual, and there is no mechanism here that allows for this knowledge to spread through the population and become global. This is the situation in which everybody knows some bars and goes to one of those again and again without paying attention to whether other people, strangers or acquaintances, also go to those bars. In order to see structure emerge in the environment, so that some patches are more visited than others, knowledge about options must spread through the population and become correlated among individuals. For the knowledge to spread we need some form of information sharing between agents.

3.2 Model 2 — Agents listening to others

In the previous model, agents acting independently on the basis of their own individual experience, choosing to go to locations that they personally recognize by having visited before, did not suffice to create emergent environment structure (though it could still be possible, perhaps given long enough). It seems more likely, and more realistic, that transmission of information between agents will enhance any clustering of choices in the space of options (here, locations)—the "social computation" enabled by a communicating population of simple decision-making agents should lead to greater environmental impact (as has been found in simulations where the interactions of many generation of simple language learners enable syntax to emerge—see Kirby, 2001). This information transmission can be accomplished either directly, through communication in which agents to tell each other about locations that they recognize, or indirectly, through agents observing the actions of others. We first consider the former situation before turning to models of indirect communication in the next sections. To add direct communications to our models, we must specify who can talk to whom, how often, and about what.

In earlier models (Todd & Kirby, 2001) we found that when individuals could hear from one other agent at the same location (that is, agents who have currently made the same option choice) about one location that the agent recognized, this could foster the emergence of clustered or J-shaped distributions of agents over choices. Such environment structure did not emerge particularly readily, though; if individuals told others about *any* randomly-chosen location they currently recognized, there was no effect. This was because many of the locations that a given individual recognized were known because they had been heard about from others, who may also have heard about them from others, which means it could have

been a long time since any of the agents in this communication-chain had actually personally been to (chosen) that location. This time delay meant that the agents' recognition knowledge could be out of step with the actual choices currently being made by others in the population (also indicated by a low choice/recognition knowledge correlation), keeping choice clusters from appearing. When we restricted individuals to talk only about locations they had actually been to recently, and thus recognized from personal choice, the temporal lag in communication was reduced, and agents did indeed begin to cluster more strongly on particular locations.

In our new models, we relax and simplify the communication somewhat. Instead of only listening to other agents who have made the same current choice (are on the same location), now individuals can hear from all the agents in the population. And instead of announcing a location that they recognize (at all, or only recently), agents now mention just the location they have currently chosen. Thus, on each time step each individual hears from one randomly-selected other agent about the location that that agent is currently on. When one location happens to have more visitors than average, it will be heard about more than average, and stored in the recognition memory of a number of individuals, influencing their later choices. What happens when this form of direct communication is used along with the recognition and fluency heuristics?

When we look at the distribution of agents across locations in the last 100 time steps of a 20,000 time step run, we see in Figure 2a that the recognition heuristic does not create population clusters any more than the random distribution of agents did (as shown in Figure 1b). However, when agents can use continuous recognition memory to distinguish between locations they have been to and possibly heard about more often or more recently, a strong J-shaped distribution does emerge (Figure 2b). This means that directly learning about a location where there are currently other agents (or in other words, hearing about an option that others have currently chosen for themselves), particularly when more popular options are more likely to be learned about, can allow agents to coordinate their knowledge and their choices sufficiently to produce a degree of conformity and thereby shape their environment appreciably. This is certainly what we would generally expect from observing cultural conformity in the real world, where people do talk about their choices with each other all the time; the interesting aspect of this model is that this structure can appear just through the use of so simple a choice mechanism that just relies on recognition knowledge.

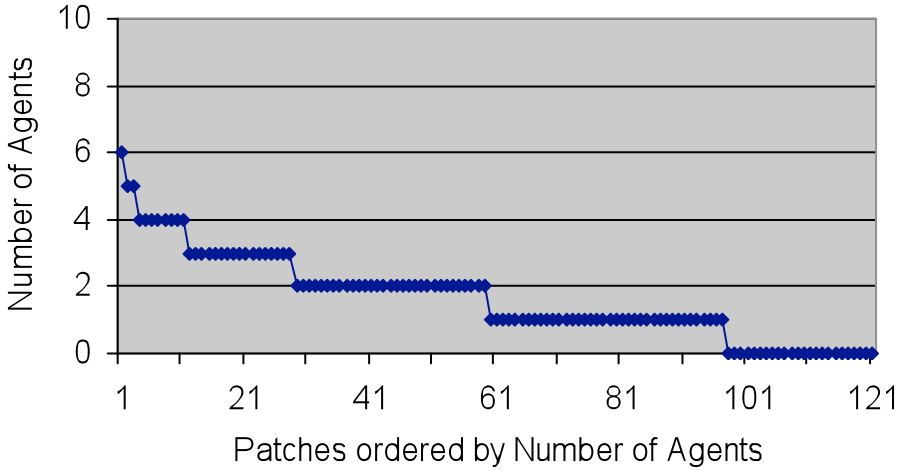

Figure 2a. The distribution of agents among the patches averaged over the last 100 time steps of a 20,000 time step run, with patches rank-ordered from left to right by number of agents present. Agents used the binary recognition heuristic, and no clumpy patch structure emerged.

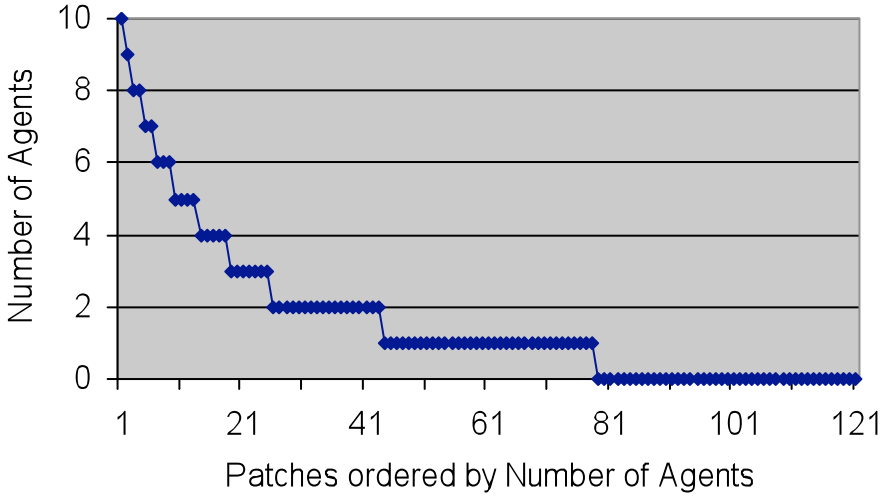

Figure 2b. The distribution of agents among the patches averaged over the last 100 time steps of a 20,000 time step run, with patches rank-ordered from left to right by number of agents present. Agents used continuous recognition in the fluency heuristic, allowing a clumpy J-shaped distribution to emerge.

3.3 Model 3 — Agents counting others

Direct sharing of information between agents allows them to coordinate in such a fashion that general agreement and conformity of choices develops at the population level. But is this direct communication a necessary component for such environment structure to emerge? What happens when agents can only indirectly influence each other's choices? When agents "share" their knowledge and behavior by simply paying attention to how many other agents have made the same choice and strengthening their recognition memory according to this count, this proves to be enough to allow coordination once again. As shown in Figure 3, both the recognition and fluency heuristics lead to some locations becoming known by all the agents in the population, which in turn creates a J-shaped distribution of agent choices and a high choice/recognition correlation (.44 and .75 respectively).

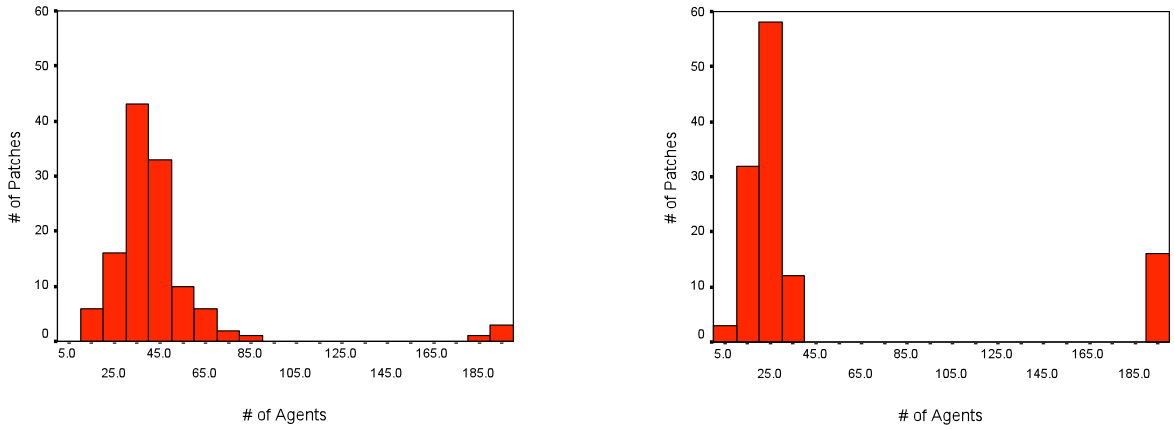

Figure 3. Histograms of the number of patches (y-axis) that are known by a certain number of agents (x-axis) after 20000 time steps for runs with agents using the recognition heuristic (left) and the fluency heuristic (right) with a weight-given-to-other-agents of 1.0 and a memory-trace-decay-rate of 0.1.

The structure found in the environment stems from an inequality in how well the patches are known by the agents. To understand how this inequality arises, it is important to remember how the social information sharing rule of this model works. When an agent goes to a certain patch, it stores this patch in its recognition memory with a value of 1.0, plus an extra value of 1.0 for every other agent that is on that patch at the same time. When many agents are on a certain patch, these agents all store the patch with a high value in their memory. Those agents now recognize this patch for quite some time. Since agents—if they can—go to patches they recognize, it is likely that those agents end up returning to that patch

again in the future. Agents arriving later on that patch for the first time will probably meet more than an average number of agents there, and will also remember the patch well. Thus as soon as a patch has many visitors, which early on may happen accidentally, a self-reinforcing mechanism is kicked off that eventually can lead to the situation in which that patch is known by all the agents.

There are more locations known by the entire population when continuous-valued recognition is used (with the fluency heuristic) than with binary-valued recognition for the memory-decay settings used here. This can arise because the fluency heuristic allows more discriminating choices between options than does the recognition heuristic—the former allows the most recognized of all of the recognized options in a choice set to be chosen, while the latter leads to random selection from those recognized options. Thus the fluency heuristic enables agents to return to more-recognized locations preferentially and thus to build up even more recognition of those locations. This in turn means more chance to return to that location again in the future, and hence more agents at that location at any point in time, which also leads any other agent who ends up visiting that location to note the increased number of others and hence to store that location strongly so that it too is likely to return there. In this way, more locations will become known by all agents more quickly than if they used binary recognition. However, when the presence of other agents making the same choice is given greater influence (that is, the count of other agents increases the strength of the recognition memory activation even more), this pattern can reverse, with fewer locations known by everyone in the continuous case than in the binary case. The reason for this reversal is that greater influence can cause some locations to become widely known even more quickly, and once this happens, the agents are likely to choose to go to that small set of locations exclusively (if given the choice), so that those initially-popular locations alone become more and more visited and known. In other words, the stronger feedback process created by greater social influence can result in rapid convergence onto a smaller set of options, effectively shutting out the competition.

3.4 Model 4 — Agents recognizing others

In this last model, the simulated individuals are a bit more discriminating. They no longer pay attention to every stranger, but instead only pay attention to their acquaintances. This means that instead of counting the number of other agents on their current patch, agents just count the number of agents they *recognize* that are on the patch. Under these circumstances

different patterns arise, depending on how strongly agents attend to the presence of their "friends".

When individuals only pay little attention to the other agents they recognize, and when they forget about patches and agents rather quickly (after ten time steps), no environment structure emerges. In this model every agent an individual meets is stored or updated in the individual's memory with the standard value of 1.0, and this memory trace decays every time step (the same as for the location memory). Looking at that memory for agents, it can be seen that the individuals in this model only know (recognize) about 16 other agents on average. Since at the beginning of each run no patches are more known than any others and all agents are equally likely to be on any of the 121 patches, the chance that an individual meets a friend (recognized agent) again before he forgets about him is consequently very small. So while agents in this model do pay attention to other agents they know, since they almost never meet again there is effectively no influence of this agent recognition, and so no structure will emerge.

What if agents were more impressed by seeing someone they recognize? When we increase the weight of attention paid to every other agent to 4 (rather than 1), friends will now be remembered four times as long, and any patch where two friends meet is stored with an extra value of 4 in both their memories. Now again we see structure emerge, with fewer universally-known locations when continuous recognition is used, again because of the feedback processes operating. In that fluency-use case, because there are fewer patches that are very well known and thus well visited, agents are likelier to meet, simply because there are fewer "meeting places" and so the population is less spread out. On the other hand, agents using binary recognition might select and end up at the less well-known patch from the choice set because they do not use how well they recognize a patch which could also indicate how well a patch is known. In the latter situation, when agents end up at the less known patch, they are also less likely to meet many other agents there, which means they will store that location less strongly in memory, and return to it with lower probability, making that patch less likely to become widely known.

What can be concluded from these models is that the emergence of environment structure, in the form of J-shaped distributions of agents across chosen options and universal recognition of a few options, is enhanced by a slower memory decay rate for recognized patches, a slower memory decay rate for recognized agents, and a greater weight or influence given to the presence of other agents on the same patch when storing patch recognition.

Furthermore, how that influence of other agents is distributed makes a difference: When much attention is paid to a small group of agents (e.g., only those recognized), structures emerge less easily then when less influence is spread out across more agents, even when the total amount of influence is made equal. Furthermore, inequality in how well patches are known is not a guarantee for structure to emerge in the environment in terms of choices made (i.e., distribution of agents across options). In the models using binary recognition, where agents do not go to patches they know best, there must be a large difference in how well patches are known for choice structure to show up in the environment. However, when there is inequality in how well patches are known when continuous fluency is used, these knowledge differences will almost immediately influence the choice structure, because of the stronger feedback loop enabled by the more discriminating fluency-based decisions.

4 What we have learned, and where to next

Agreement can be useful. Even when there is no independent advantage of choosing one option or course of action over another, it can still be advantageous if most people settle on the same option. Individuals can share the knowledge they gain about this common option with others (e.g., how to fix the latest wormhole in a Microsoft product), allowing them to get more use out of it. Individuals can coordinate with each other for different purposes through selecting the common option (e.g., planning a spontaneous weekend trip with friends after meeting up by chance at the favorite hot-chocolate bar). And social cohesion can increase from the shared knowledge about the common option (e.g., more conversations around the water cooler after everyone watched the same episode of "Iron Chef"). If everyone made their own independent choice—if conformity disappeared—these social advantages would be greatly reduced.

What we have demonstrated in this paper is that it does not take much cognitive machinery to make decisions that will have a conformity-producing impact on the environment. Just using recognition knowledge, whether and how often or recently particular options have been encountered, to distinguish and choose between available options is enough to enable clustered choices to emerge—provided that the recognition knowledge is at least partly coordinated between individuals. This coordination can come about either through direct communication, in which individuals tell each other about options they recognize, or through indirect observation, in which individuals store how many others they have seen making a particular choice. And while having a more precise memory of experienced options,

in the form of continuous rather than binary recognition, helped speed the emergence of environment structure, adding extra information in the form of a memory for other agents (model 4) did not strengthen this effect.

Several other factors should be examined in more detail to fill out this story. First, we need to explore the impact of environment size, in terms of the number of options available for individuals to choose among. In the models presented here we saw two major different types of structure emerge, one in which all options were known (recognized) by some medium number of agents, and another in which some patches ended up being known by all agents. Could a larger environment lead to the emergence of multiple clusters of locations that are highly known and chosen by separate subsets of agents, as we see for instance in consumers split into different brand-loyal clans?

Relatedly, how stable are the clustered choices that agents make in these models? That is, will a group of agents who have all converged on one option dissipate over time, and be replaced by another group of similar size clustered on another option? So far, we have not analyzed *where* the clusters are, only the degree of clustering (in part because all clusters have the same quality up until now), but we expect that in the situations where only a few locations are known by nearly all of the agents, these clusters will be very stable for long periods. Patterns of chosen-option change over time in these models need to be related to similar patterns observed among consumer choices, for instance.

Another interesting avenue to explore is to make the nature of communication more realistic in our simulations. In the cases we have investigated here, each individual has an equal chance of serving as a model, or communicating, to all other individuals in the population. In reality, some models are more influential or prestigious than others (leading to a prestige bias—see Henrich & McElreath, 2003), and some objects or ideas dominate media channels. Such effects may create biased recognition without initially biased frequencies of the objects or ideas in the population. For example, teachers, political leaders, and celebrities can potentially spread recognition of their opinions and behaviors to many individuals (whether for good or ill). Cavalli-Sforza and Feldman (1981) term this one-to-many or few-to-many transmission, which they argue sometimes has important consequences for the evolution of knowledge.

One-to-many transmission processes, whatever the details of their functioning, can alter the frequencies of behaviors, objects, and ideas by creating greater recognition for a smaller number of behaviors, objects, and ideas. Essentially, some people or organizations

drown out other potential models. A recognition-based mechanism would then use this familiarity to make choices about what to imitate, consume, or trust. In this way, one-to-many transmission may lead some things to come to be both more recognized and more chosen more quickly, even when the initial frequencies and underlying qualities of the items are similar.

We would also like to make the influence of other agents more psychologically plausible, for instance by having a decreasing impact of greater numbers of other agents, as Asch (1956) found. Even more importantly, we need to include a consideration of network structure in our analyses. We have begun looking at how this environment can be considered a bipartite network with links between agents on the one hand and locations on the other. Can such an analysis help us understand how knowledge spreads through the population in different environment structures?

We have shown here that individuals can use simple decision mechanisms based on innate recognition abilities, along with direct or indirect sharing of knowledge, to link their behaviors in a way that strongly impacts the environment. Just deciding where to get your next cup of hot chocolate based on what bars you recognize, have heard of, or have observed strangers and acquaintances sipping in on your last visit, can suffice to get everyone coordinated in a world-shaping way.

References

Arthur, W. B. (1988). Self-reinforcing mechanisms in economics. In P. W. Anderson, K. J. Arrow and D. Pines (eds.), *The Economy as an Evolving Complex System*. Addison Wesley.

Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs 70*, Whole No. 16.

Boyd, R., and Richerdson, P.J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.

Cavalli-Sforza, L.L, and Feldman, M. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton University Press.

Durlauf, S. N. (1997). Statistical mechanics approaches to socioeconomic behavior. In W. B. Arthur, S. N. Durlauf, and D. A. Lane (eds.), *The Economy as an Evolving Complex System II*. Addison-Wesley.

Galef, B. G. Jr. (1996). Social Enhancement of Food Preferences in Norway Rats: A Brief

Review. In C. M. Heyes and B. G. Galef, Jr. (eds.), *Social Learning in Animals: The Roots of Culture*. Academic Press.

Gigerenzer, G. (2001). The adaptive toolbox. In Gigerenzer, G. & Selten, R. (eds.), *Bounded rationality: The adaptive toolbox*. MIT Press.

Gigerenzer, G., Todd, P.M., and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Goldstein, D.G.; Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P.M. Todd, and the ABC Research Group (eds.), *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83.

Grönlund, A., & Holme, P. (in press). A network-based threshold model for the spreading of fads in society and markets. *Advances in Complex Systems*. (Also available as an arXiv preprint at http://arxiv.org/abs/physics/0505050; downloaded 10/5/05.)

Henrich, J. and Boyd, R. (1998). The evolution of conformist transmission and between-group differences. *Evolution and Human Behavior, 19*.

Henrich, J., and McElreath, R. (2003) The Evolution of Cultural Evolution. *Evolutionary Anthropology*, 12 (3).

Hertwig, R., Hoffrage, U., and Martignon, L. (1999). Quick Estimation: Letting the environment do the work. In G. Gigerenzer, P.M. Todd, and the ABC Research Group (eds.), *Simple Heuristics That Make Us Smart*. Oxford University Press.

Heuvelink, A. (2004). *Monkey See, Monkey Do...: Modeling the formation of social consensus*. Unpublished Master's thesis, Department of Cognitive Artificial Intelligence, Utrecht University.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2).

Lane, D. A. (1997). Is what is good for each best for all? Learning from others in the information contagion model. In W. B. Arthur, S. N. Durlauf, and D. A. Lane (eds.), *The Economy as an Evolving Complex System II*. Addison-Wesley.

Latané, B. (1981). The psychology of social impact. *American Psychologist 36*.

Milgram, S., Bickman, L., and Berkowitz, O. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology, 13*.

Noble, J., and Todd, P.M. (2002). Imitation or something simpler? Modeling simple mechanisms for social information processing. In K. Dautenhahn and C.L. Nehaniv (Eds.), *Imitation in animals and artifacts*. MIT Press.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

Reimer, T., & Katsikopoulos, KV (2004). The use of recognition in group decision-making. *Cognitive Science,* 28.

Schooler, L.J., & Hertwig, R. (in press). How forgetting aids heuristic inference. *Psychological Review*.

Todd, P.M., and Kirby, S. (2001). I like what I know: How recognition-based decisions can structure the environment. In J. Kelemen and P. Sosík (eds.), *Advances in Artificial Life: 6th European Conference Proceedings (ECAL 2001)*. Springer-Verlag.