# The Role of Stability in Cultural Evolution: Innovation and Conformity in Implicit Knowledge Discovery

Joanna J. Bryson

University of Bath, Bath BA2 7AY, United Kingdom email: j.j.bryson@bath.ac.uk

**Abstract.** One of the defining traits of humanity is our capacity for accumulating innovations. While many authors focus on the innovation process itself, Evolutionary Anthropology has become more interested in the *accumulation* part of this uniqueness, and in particularly whether something like an evolutionary account of cultural acquisition can explain it. In this chapter I discuss the role and sources of innovation in generating culture, and also the role of norms in preserving it. I demonstrate through two sets of simulation experiments a model of cultural evolution exploring the problem of cultural stability and change. The first models the impact of noisy transmission and modularity on cultural stability. The second looks at the impact on cultural change if a biologically-advantageous variant emerges of a single cultural trait.

## 1 Introduction

Innovation is a topic of great interest in the study of cultural evolution [32]. How do new behaviours and ideas come to be established in a culture? The reason for this interest is obvious — culture is after all an amalgamation of past innovations, so the study of innovation is also the study of the origins of culture. However, the emphasis on novelty that the term 'innovation' elicits may not be the most useful perspective for truly understanding culture origins. For evolution, the greatest challenge is *preserving* useful traits. The most essential characteristic of life is its capacity to reproduce — diversity and even increasing complexity, while also fascinating, occur in other materials as well.

How difficult is preserving culture? One indication of this may be the fact that so few species do to any measurable extent [42]. While we know that many species use culture as a part of their adaptive repertoire, there is little evidence for accumulation across generations [48]. For some this has been seen as clear evidence that human culture is not so much a result of unsupervised processes akin to Darwinian evolution, but rather that it requires explicit and special mechanisms for transmission [17]. Of course, this argument is either creationist or cyclical. While the current importance of culture to human survival means that any innovation for preserving innovation is highly adaptive biologically as well as culturally, it is unlikely that the initial innovation that supported cultural accumulation was acquired culturally [12, 13].

Sperber and Hirschfeld [37, 38] argue that due to the noise inherent in the social transmission of behaviour, only a modular model of learning and mind can explain cultural preservation and stability. They propose that the massive modularity hypothesis

(MMH) [16, 36] is an alternative to the current emphasis by many evolutionary anthropologists and behavioural ecologists on imitation as the source of culture [29]. Unlike the modularity of Fodor [24], under the MMH modules are acquired during development from culture. For example, the simple fact that a unique word exists is a clue for the related existence of a useful concept, and a child will search for a robust application of a novel term [44]. This can apply similarly to more complex cultural input such as stereotypes [18] or myths [39].

In this chapter I demonstrate through simulation the robustness of imitation learning given the context of a modular culture. This is simply because errors in transmission tend to cancel each other out provided there are sufficient exemplars and there is no bias in their production. I then go on to show however that a bias towards conformity is absolutely essential for preserving innovation, and in fact demonstrate the difficulty of constructing a model that can discover optimal new behaviours. At least for the simple abstract simulations presented here, it is difficult to find a single set of parameter values that both allow for innovation and preserve good solutions once found. I suggest that this may explain why cultural species do not tend to have a single such set of parameters, but rather to be neophilic and neophobic at different ages or in different contexts.

#### 1.1 Terms and Concepts: Cultural Evolution and Innovation

Whether culture can be usefully thought of as an evolutionary system is a matter of ongoing debate [1, 35, 51]. While few doubt that the biological capacity for culture must have evolved, the question is whether or not culture *itself* evolves. That is, are the contents of cultures themselves (e.g. words, ideas, symbols, images or even just socially-aquired behaviour) subject to reproduction, variation and selection in a way that is meaningfully similar to the process Darwin identified as explaining the origin of species. While acknowledging this controversy, in the present chapter I will not address it directly, but rather just assume an evolutionary perspective towards culture. This is a standard approach for simulation: to the extent that any results are validated by comparison to their target system in the natural world, these results can be seen as also verify the initial axiomatic assumptions behind the simulation as well [14].

Taking then a selectionist perspective, we might usefully view innovation as errors in the cultural replication and preservation process that happen to persist. Of course this perspective is a simplification. There may well be intelligent search performed by some individual 'carrier' of the culture that is the root cause of some specific 'defect in replication'. Cultural evolution is not *necessarily* an unsupervised and unintentional search process. Further, there is no reason for inheritance in cultural evolution to be limited to one or two parents and a single recombination event [11, p. 89–90]. Rather, the more information that can be gathered, the easier it is to detect the salient signal inside the noise and irrelevant detail. Finally, any particular variation in culture may actually convey a *biologically*-adaptive benefit, so it may not just 'happen' to persist. However, taking a simplified meme's-eye view of innovation may help us understand the processes that underly cultural change [20].

I take it as given that *some* cultural variation happens as a result of blind chance and copying errors. For the sake of simplicity therefore, this will be the only sort of 'invention' in the models described here. Presumably intelligent invention only accelerates the

pace of change by making actually adaptive 'errors' more frequent, but otherwise does not substantially alter the process. In an effort to keep this chapter as clear as possible, I will call any deviation from a previously-established culture an invention, and any invention that reliably persists through cultural transmission an innovation. The models below show conditions where an adaptive innovation can be made, and conditions where innovations occur even though they have no adaptive impact.

## 2 Background: Modularity and Cultural Stability

It is useful to decompose the social communication of behaviour into two different levels. For the purpose of this chapter, I will refer to the rote replication of end effector positions or end effects as *imitation*. This is a simplistic, ordinary-language use of the term, but sufficient for the experiments described here<sup>1</sup>. By 'imitation' I do *not* necessarily mean a full transfer of behaviour. This latter would imply that two agents have communicated not only actions but a model between them, such that they have the same understanding of the role of the actions they imitate, and the goals they might meet with those actions. Sperber and Hirschfeld argue that this shallow sort of imitation cannot be an integral part of cultural transmission. Although I generally find their work very useful, my main departure is that I believe shallow imitation *can* fulfill this role.

Sperber and Hirschfeld [37, 38] argue that due to the unreliability of both performing actions and perceiving others' acts, reliable cultural transmission is exceedingly unlikely. Giving evidence based on the known degradation of signal experienced in simple transmission chains of spoken sentences (e.g. the party game of Telegraph [USA] or Chinese Whispers [UK]), they criticise the current emphasis on the role of imitation in cultural transmission. Imitation is limited to mere replication of apparent behaviour, and that is in turn limited by constraints in our ability to perceive others' actions, and also by our own inability to execute our actions exactly was we intend. Sperber and Hirschfeld insist that what matters is the deep transfer of mental models from one mind to another, not the shallow imitation of expressed behaviour.

How can this deep model be recovered from limited perceptual information? Sperber and Hirschfeld see no way, and use this implausibility as evidence that some information must come from elsewhere. They suggest this missing information is the information encapsulated in modules. Modules under the MMH may have both genetic and explicitly-learned components. Thus extra information is available to compliment the shallow information available from perception and imitation.

People used to implementing artificial learning systems and / or familiar with the mathematics or logic of learning may find the above arguments somewhat unsatisfying. After all, provided that errors in perception or action are random, they can be considered noise and will cancel each other out provided there are enough observations. Anything left is not random is also not noise, but rather some sort of signal which ought to be useful. However, this only accounts for part of the Sperber and Hirschfeld argument, and the other part (that some information is missing) I think is correct, though their theory is under specified. Where does the extra information they postulated as coming from

<sup>&</sup>lt;sup>1</sup> For more elaborate definitions see e.g. Bryson [12], Whiten and Ham [47].

modules *itself* originally come from? Biological evolution, cultural evolution and individual learning are all forms of learning. Therefore taken as sources of information and knowledge, their power is essentially identical [4, 53]. Thus to some extent the Sperber and Hirschfeld argument is overly compartmentalised. To say that the extra information required to make sense of the noisy social transmissions *comes from* modules is still to beg a question of how the modules themselves have come to support this process.

Although they are not completely explicit about it — in fact, they are almost explicitly agnostic on the topic [38, p. 41] — it seems likely Sperber and Hirschfeld are implying that some of what we commonly call 'human culture' is genetically encoded. This is problematic if we take the simple information-centred definition of culture I ordinarily favour: that culture is all behaviour acquired from conspecifics by non-genetic means [12, 35]. However, taking instead a more ordinary-language view of culture as the aspects of behaviour such as language and social organisation which seem to vary between peoples, then the idea of a genetic component becomes more sensible. There is relatively little controversy for example that *some* aspects of linguistic competence must be genetic (such as the capacity for vocal imitation and transmission), though others are clearly learned by individuals from their own or another culture [23]. Given what we understand about how learning affects evolution [2, 6, 27], we should also expect that some things that may first evolve as cultural variation could over time become at least partially genetically entrenched.

#### 2.1 Modularity and Learning

What Sperber and Hirschfeld really propose then is that the automatic or implicit learning of culture from imitation cannot in itself account for all the richness of human culture. Although they acknowledge a possible complementary role for imitation-driven cultural transmission, their own emphasis is on complex mental models underpinning human behaviour. This process in turn requires the explicit transfer of abstract / symbolic knowledge. Symbols in themselves contain almost no information, but cultural participants who understand them have high-information-content associations, or *grounding*, for them. Under the Sperber and Hirschfeld model, grounding is encoded in modules and contains most of the information necessary for the newly acquired behaviour.

This notion of the role of modules is quite similar to one I have proposed in the context of artificial intelligence [9, 10]. In this work I extended the model of modular organisation of intelligence known as Behavior Based Artificial Intelligence (BBAI) [8] to include module-based learning. The original insight of BBAI was that real-time intelligence is best decomposed into behaviour modules. 'Best' in this context means

- responsive to the demands of an unpredictable and rapidly changing environment,
- robust to the difficulties of both sensing and control, and
- easily and reliably developed by programmers and roboticists.

Under standard BBAI, the purpose of a behaviour module is to perform some action or provide some capacity for its host agent. Modules consist therefore of instructions for whatever control is necessary for those actions, but also of whatever perception is necessary to guide those actions. This tight coupling of sensing to action is a hallmark of BBAI. It simplifies the problem of building intelligence by restricting the problems worked on to a minimum set of capacities each with only the most essential detail required to reliably execute its tasks. The strength of the approach was not only argued but also demonstrated in the first robots able to move autonomously at animal-like speeds [7, 28].

The Bryson [10] extension to BBAI stems from the observation that perception is more than just sensing. At any one instant, sensing provides too little information to successfully disambiguate the correct next action. Animals address this problem through systems of memory ranging from integrating recent signals through conventional ideas of memory (e.g. map learning) and on through genetically provided biases [15, 34]. This applies to BBAI as well. Just as behaviour modules should contain the dedicated and specialised sensing necessary for their actions, they should also contain the dedicated and specialised memory necessary for both perception and control. One advantage of this modularisation of learning is that specialised representations can be chosen that facilitate the particular sort of learning that each module needs. This increases the probability that the individual agent will learn and act successfully [52].

### 2.2 Bootstrapping Culture: The Law of Large Numbers

From the above review it should be obvious that I strongly support the idea that modules can and almost must support all learning. Strictly speaking, modular learning system cans generally be recast as a computationally-equivalent homogenous ones. That is in theory a homogeneous system can learn anything a modular one can [52]. However, accurate learning is much, much less probable without bias which modularity can provide, and therefore will take much longer on average to converge [5]. For an animal or other real-time system, this means it is less likely to succeed in time to be used.

This result includes the individual learning that underlies cultural transmission and evolution. However, we must consider the full process of internalising information to guide behaviour, from evolution through development and learning. We also need to account for cultural transmission in the non-human species in which it has been observed [22, 30, 33, 43, 46, 50]. Even ants might be thought of as having minor cultural differences between colonies, since their members both determine and learn new nest locations in a distributed, social manner [25].

Sperber and Hirschfeld are correct to be skeptical of one-shot imitation as a mechanism of social transmission. Essentially, if a single signal can transmit enough knowledge to really alter behaviour, then that knowledge must have been previously accumulated and stored in such a way that the behaviour observed has information-equivalence to a symbol anyway [54]. In this case, imitation is not fundamentally different from explicit communication. There will in fact be a continuum of conditions whereby true communication of cultural contents can be achieved with more or less information prompted, depending on how much the cultural and genetic predispositions of the demonstrator and the receiver align. To return to the telegraph metaphor, the way real telegraphs work is through a system of repeaters that can remove noise accumulated and re-boost the signal. Where the repeating process is intelligent, degradation is probably even less of a problem.



Fig. 1. Culture degrading. Notice the presence of subcultures among neighbouring adults.

## 3 Experiment 1: Stability of Culture with Noisy Transmission

The following experiments demonstrate the above arguments, and then move to explore some of their consequences. They are abstract and I have not been able to think of a good way to validate them, so at this stage of development they should probably be thought of as no more than intuition pumps [21]. Here I present a modular model of a culture. The model is agent-based (ABM). It is built in NetLogo, a standard and freely-available ABM development environment [49]. The code for the model is available from the author by request, or her Web site by demand.

#### 3.1 Model

An ABM consists of three parts [14]:

- 1. an *environment* where the agents are situated and which determines their possible behaviour;
- 2. *attributes*, also known as parameters or variable state, which describe the agents and what makes them individual; and
- 3. behaviour or intelligence, the actual algorithms which the agents use for control.

I describe each of these in turn.



**Fig. 2.** Culture recovering. The probability of generating incorrect actions has been reduced just 4%.

**Environment** The first model has a very simple environment. It is entirely social, with no intrinsic reward provided for any behaviour. Space is described as a torus — that is, a square with the left and right edges connected, and also the top and bottom ones. This means that the code and analysis do not have to deal with exceptional agents that live at the edge of their world. Agents occupy every possible location in the grid; each has eight neighbours it can observe.

Agent Attributes Agents have three types of attributes [14]:

- 1. static parameters which vary only between experimental conditions,
- 2. *run-dependent parameters* which vary per run and often per individual but are fixed at the beginning of the run, and
- 3. dynamic parameters which change within a single agent's lifetime.

Besides having eight neighbours, the most fundamental static parameter in this model is the agents' modules. All agents have the same number of modules. Although the exact number of modules is run-dependent, how they operate is static. Each module is very simple — it essentially corresponds to a context the agent may find itself in. Each agent has a single behaviour that it currently expresses in that context; *which* behaviour among many possible is learned socially (see algorithm below). For convenience in visualisation (but not in explication) there are exactly as many possible behaviours for

each context / module as there are modules. It would probably be sensible in future versions to allow two different N for modules and possible behaviours for greater clarity in communicating the results.

Since the agents acquire their behaviour socially, they need to be able to keep track of other agents' behaviour they witness. Thus each agent has associated with each module a memory. The size of this memory is the same as the number of possible actions. The agent remembers how many times it has seen each action it has witnessed in each context. Thus the content of this memory is a dynamic parameter.

Besides the contents of its memory, the only other dynamic parameter of an agent is its age. At the very beginning of a simulation, age is assigned randomly to each agent from the full range of possible values. Subsequently, any new agent starts with age 0.

In addition to the number of modules, there are a number of other run-dependent parameters:

- Each agent's (X, Y) position in social space. This determines which other eight agents are its neighbours.
- The number of 'years' spent as a child and as an adult. The difference is that no one learns socially from children.
- The number of acts performed per 'year'. This in combination with the lifespan and the size of the culture determines how much each agent will experience in its 'life'.
- The probability of a perception error and the probability of an action error. If one agent performs an action error, all of its neighbours will see an unintended behaviour in a particular context. If one agent experiences a perception error, then it is the only agent that's knowledge is affected. In both cases, an error means a value for an action is randomly drawn from all possible acts. For the sake of simplicity, in the experiments discussed here the only probability varied was of action error. This is more likely than perception error to cause perturbations of culture, since it can bias eight neighbouring agents" beliefs in the same way.

This variable is somewhat dynamic, in that it can be varied during the course of a simulation by the experimenter. This allows for a relatively easy search for a threshold value below which the culture is stable an above which the culture degrades. However, nothing the agents do themselves changes this value, so from their perspective it is run-dependent.

- The weight given to the seed culture at the beginning of the simulation. At the beginning of the simulation, all of the first generation of agents have their memories set to some initial cultural value for each context. This value is set by the experimenter. If the weight is five, the agents have a memory equivalent to having seen other agents perform that action five times. This parameter has no other role in the simulation after the first generation has died.

For visualisation, the field of agents is visible as a square. The agents are arrow shaped. The agents are coloured to indicate their age: children are light and adults dark. The viewer can be set to examine any one behaviour context for all the agents. The beliefs and therefore the chosen action of each agent for that context is then visualised as the angle at which the agent points. The angle = (360 \* i)/N, where i is the number of this particular context, and N is the number of contexts and therefore also the number of possible beliefs. As a secondary visualisation, there is also a chart which shows the

percentage of agents that conform to their original beliefs in the seed culture for the first four contexts. Since all contexts are functionally identical, these first four can be treated as a small random sample of modules.

Agent Behaviour On every program cycle, a context is chosen by the environment at random. Each agent then checks its memory for that context and expresses whatever action it has itself most often witnessed in that context. If more than one action is tied for having been witnessed the maximum number of times, then the tied actions are chosen between at random. Assuming there is some Probability of Action Error (PAE), the agent then has a PAE chance of choosing an action randomly from all possible values and expressing it. Otherwise, it expresses its module's true value.

"Expressing an action" in the simulation is manifest as an agent asking all eight of its neighbours to add one count to that action's value in that context, indicating that action / context pairing has been witnessed once more. If there were a probability of perception error, at this point a random value might be introduced into an individual's memory rather than the act expressed. However it is best practice to limit the number of parameters on a model for simplifying analysis, and for the reason stated above I chose only to manipulate action errors for the experiments presented here.

When an agent reaches its age limit, it dies. When an agent dies, it is immediately replaced with a new agent of 0 age. This new agent has a completely empty mind. It has the same number of modules as the rest of the agents in the simulation, but every possible value for every module is given 0 weight. Thus its initial actions will be entirely random.

### 3.2 Results

Cultural stability is directly correlated to the number of exposures to an action that an agent is likely to experience for each action in its lifetime. Thus the longer adult life, and the more actions that occur per year, the more stable culture. On the other hand, having more modules decreases the number of actions *per* module, so this is negatively correlated to stability, as of course is the PAE.

One surprising result concerns the influence of children on culture. The tendency to ignore children's behaviour (which is initially essentially arbitrary) has been proposed as a mechanism of cultural stability. However, because even children after one year are more likely to express their culture's values for any module than any other value, short-ening "childhood" — or at least, the period where children do not serve as cultural models — actually *increases* cultural stability. Of course this is not the only attribute of childhood. If I had modelled it also as a period when more is time devoted to observation of others (perhaps by increasing the neighbourhood size for children), then a longer childhood might have been more beneficial.

Figure 1 shows a run with parameters set such that the culture is fairly stable, but not sufficiently so to stop degradation (forgetting) of the culture. Since we are observing the i = 0 context module, the agents conforming to the original culture are pointing straight up. Notice that young agents (the light / yellow agents) may be oriented in any direction since they will not have seen many expressions of behaviour in this context yet. However, where adults (dark / blue agents) are misoriented, they often are so in company. Thus the same mechanisms that largely preserve culture can also serve to form and preserve subcultures.

Figure 2 shows the same simulation in the future. However, just after the previous snapshot, the probability of action error was lowered from 94% to 90%. Notice this does not simply freeze the decline of the culture, but actually results in the initiation of a rapid recovery. This is because the level of conformity to the original culture was still > 1/N. If culture had degraded to total chaos, then reducing the *PAE* would have lead to conformity as well, but not necessarily to the original value. Note also that a culture will never have 100% conformity because of the ignorance of children, but with a low *PAE* a stable culture will achieve a high level of conformance.

#### 3.3 Discussion

The idea that a module might take only a few discrete values may seem such an extreme abstraction that it renders the model meaningless. However, we know that animals including humans are extremely inclined to categorise perceptual data. Even in continuous domains such as the light spectrum, humans are far more sensitive to variation near the "boundaries" between named colours than well within them [26, 40]. This emphasises the role both Sperber and Hirschfeld and I hypothesise for modules in learning in general, of which social learning is a special case. Through some combination of genetics and experience the agent is assumed to know a set of categories or concepts, which learning facilitates a choice between.

Social learning may also facilitate the discovery of *new* categories and modules by signalling through variations in behaviour a perceptual difference an agent had not otherwise detected [3, 11]. However, module construction is not modelled in the current simulations.

## **4** Experiment 2: Innovation

In the first model we already witnessed the formation of subcultures. Since these can be stable for a few years or even generations, they might already be viewed as innovations. In the second set of experiments we observe what happens when one possible value for a culture model is more adaptive than the one currently dominant in the culture. To do this, we have to introduce reproductive variation into the model.

In the previous simulation, reproduction was always at exactly replacement rate. To keep the experiment simple, a mechanism of selective reproduction was chosen that kept a full environment as the *maximum* number of agents. Thus, for the non-adaptive culture values, reproduction was lowered below replacement rate.

#### 4.1 Model

The model is largely as described before, with only one exception: reproduction.



Fig. 3. The point in time when an adaptive innovation is just beginning to dominate a culture.

**Environment** The environment is largely unchanged, except that there is now one context which can be differentially rewarded. Which context this is can be set by the experimenter.

**Agent Attributes** There is one new attribute, a run-dependent parameter reflecting Selective Advantage, *SA*, described below.

Agent Behaviour One module or context is chosen by the experimenter to be selectively rewarded. For that module, only one action is right, or put another way, only one belief is 'true'. When an agent dies, if it does not hold the correct value, then its probability of being replaced is reduced by SA. On the other hand, if an agent does have the adaptive belief, not only will it certainly reproduce, but also if one of its neighbouring spaces is available, it will create one additional offspring. This allows the recovery of the population.

Note that because all agents are identical, there is no change in *genetic* distributions due to this advantage. What a parent leaves to its child or children is only its neighbours — its social network.

#### 4.2 Results and Discussion

Surprisingly (and ironically), my explorations of the parameter space have shown that a culture needs to be strongly disposed towards stability in order for a new tradition



**Fig. 4.** The impact of adaptive selection for a new value in one module on the cultural values of another. The module governing the behaviour shown has no 'true' (adaptive) belief; sub cultures for this behaviour have emerged as a result of the social dynamics resulting from selection on another behaviour.

to take root. If culture degrades easily, then even when agents stumble on the adaptive subculture they forget it again within a few generations. Obviously, however, it takes considerable disruption for a stable culture to lose its existing values so it can change to the adaptive ones. As the model is currently built, this disruption takes the form of the loss of neighbours and therefore the lower probability of discriminating the cultural values accurately. When one isolated subculture does stumble on the adaptive value and begin refilling the space around it, then the propensity for stability returns.

If the culture parameters are set to a lower level of stability, then the dominant culture can stop dominating earlier, but any new subculture has significantly more difficulty becoming self-maintaining. The adaptive subculture in particular becomes surrounded by juveniles (that is, relatively young agents) filling empty spaces. Due to less prior experience, relatively young agents are more open to influence — both to random patterns of other juveniles and to the influence of members of other neighbouring subcultures. Because it will still be disproportionately wide-spread in the culture, the ring of juveniles is particularly vulnerable to invasion by the original, non-adaptive value held by that culture. Since they surround the core of 'true' (adaptive) believers, they will generally sway their behaviour and the true belief is lost.

Another significant factor determining the outcomes for this simulation is the probability of stumbling on the correct answer in the first place. Recall that in all these simulations all behaviours are equally probable for naïve agents. If there are too many possible values for the module that is subject to selection, the agents are unlikely to find the rewarded value in time to save themselves from extinction. If the simulation were changed so that the agents were even slightly more intelligent in their search — for example, if they could remember neighbours that failed to reproduce or succeeded in having two children, this would increase the probability of the correct action being chosen.

Another interesting result is that although only one module was subject to selective pressure, the cultural norms for other modules also change. This might be because the same agents that are likely to discover the adaptive innovation had a general tendency for invention. Although all the agents have identical programs and are seeded randomly at the beginning of the simulation, the population is not entirely homogeneous. Chance patterns of distribution of age — the only differentiation between agents in the initial population — can lead to some patches of space being more or less likely to deviate from the cultural norm and form a subculture. Due to the policy of reproduction by replacement, age patterns are fairly stable. Another explanation is that change simply occurs due to the drop in cultural stability with the reduction of numbers. However, since the other modules are not having their original culture actively selected against, in some cases they recover their original value after the population stabilises (see chart in lower right of Figure 4).

Another unanticipated result from this experiment was that the pattern of regrowth after the adaptive behaviour was discovered lead to large regions of adjacent age cohorts. This in turn seems to lead to the emergence in many but not all of the module contexts not subject to selection of multiple stable cultures. Figure 4 shows an example of one such. This may have analogues in natural culture, where age cohorts may communicate predominantly internally rather than mixing with other ages. Even where there is a mix of ages, it is possible for age cohorts to focus their social learning attention on their peers.

The figures show a run where the PAE was set to what was in the non-selective condition a fairly stable value, particularly given the number of modules in the culture. Figure 3 shows the cultural values for the context and module subject to selective pressure when the number of agents holding the adaptive belief has just begun to outnumber number conforming to the original culture. Figure 4 shows the same run after the population has recovered. This second figure observes not the context subject to selection, but one of the other contexts where the values are arbitrary from a selective perspective. This context has now formed multiple sizeable, stable subcultures. Notice the pattern of ages in the agents as indicated by their colour.

## 5 Conclusions and Discussion

In this chapter I have examined and to a large extent supported the proposal of Sperber and Hirschfeld [37], while at the same time clarifying some details of how their system might work. The modules they describe utilise information previously acquired either by the species (encoded genetically) or by the individual's learning, which of course may also be channelled by the species through culture. The model I have presented demonstrates the ability of a culture to be stable in the face of enormous errors in communication. The famous 'poverty of the stimulus' is simulated by the high level of noise in the actions actually generated by the agents. Agents are nevertheless able to derive a signal because of the Law of Large Numbers and the fact the errors are unbiased. In these simulations all behaviour contexts are equally probable and all social demonstrations equally salient. In human culture we know that rare but important cultural behaviours such as rituals tend to be associated with high emotion salience indicators such as music which may assist in emphasising particular memories [31, 45]. For example, in medieval England the relatively boring and seldom-performed but essential task of patrolling the parish boundaries was made salient to young boys by beating them at boundary stones so the boys would remember the stones' locations [19].

The models also show circumstances in which innovations can not only take place but take hold. Strong tendencies towards conformity can give rise to small stable subcultures even in strictly arbitrary environments, as shown in Experiment 1. Experiment 2 explores the conditions necessary for acquisition of a newly-adaptive norm — that is, an action selected by the environment. In addition, it also shows that society-wide displacements of one cultural norm for another can take place for no direct adaptive reason, but simply as a side-effect of the disruption to the society necessary for another, more urgent change in cultural norms. This incidental disruption could be dangerous if a norm that is adaptively-neutral in the current, local environmental context actually held adaptive salience in some larger-scale environmental context, for example in times of a natural disruption such as flooding. On the other hand, if the society is too conservative — that is, makes too few 'errors' in behaviour replication, then inventions seldom occur and innovations are never adopted.

One difference between my work and that of Sperber and Hirschfeld --- I do not believe they are correct to assume that identical internal models necessarily underlie apparently identical connections between contexts and expressed actions. The conformance demonstrated here is based on shallow imitation. To some extent, it is quite likely that agents with similar brains and similar experiences will wind up forming similar internal models or theories in order to generate similar behaviour. However, it is possible that multiple models would result in the same or at least categorically indiscriminable behaviour. For example, one might obey law due to concerns about an afterlife, due to an elaborate model of the importance of the rule of law and the power of social contagion, or simply because one is evolved to unthinkingly behave like others around you, and most of them are lawful. These three models would be indiscriminable from the perspective only of your observing the law. Steels and Kaplan [41] demonstrates a robot model for this phenomenon. The underlying lexicon models for robots that have "perfectly" learned a shared language can be clearly seen to differ. In all circumstances the robots use the same terms to reference the same objects, yet the internal representation they require for grounding the terms as mappings to their sensor and motor states vary considerably between robots. Thus model conformance is not a necessary part of social conformance, and may in fact provide a useful source of variation to the populations' inventions.

The simulations I have described beg much further analysis. For example there should be a more thorough exploration of the effects of developmental differences in communication on the adaptation of cultures to new circumstances or to the opportunities of adaptive innovations. Further, the spontaneous emergence of stable subcultures in both sets of experiments might be seen as examples of sympatric speciation — a process normally attributed to sexual selection. Clearly no equivalent of sexual selection takes place here. Although the model is intended to be one of cultural evolution, it might easily be extended to model biological evolution to study this process. Or, one might hypothesise that cultural evolution underlies the beginning of sympatric speciation, and the process is then genetically consolidated. These projects are left as future work.

#### Acknowledgements

This research was inspired by an informal talk by Dan Sperber in the Spring of 2008 at the Konrad Lorenz Institute for Evolution and Cognition Research, where I was supported by the institute as the Hans Przibram Fellow with sabbatical assistance from the University of Bath. Thanks to Christophe Heintz for his discussion and comments on the original version of this paper, which was also presented to *The Fall AAAI Symposium on Adaptive Agents in Cultural Contexts (AACC '08)*, and appeared in its informal proceedings (A. Davis and J. Ludwig, eds). Thanks to that symposium also for their comments. Effort on completing the final version was sponsored by the US Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-10-1-3050.

#### References

- [1] Aunger, R., editor (2000). *Darwinizing culture: the status of memetics as a science*. Oxford University Press, Oxford.
- [2] Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30:441–451.
- [3] Bates, E. (1999). Plasticity, localization and language development. In Broman, S. and Fletcher, J. M., editors, *The changing nervous system: Neurobehavioral consequences of early brain disorders*, pages 214–253. Oxford University Press.
- [4] Best, M. L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, 7(3/4):289–306.
- [5] Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer, London.
- [6] Borenstein, E. and Krakauer, D. C. (2008). An end to endless forms: Epistasis, phenotype distribution bias, and non-uniform evolution. *PLoS Computational Biology*, 4(10):e1000202.
- [7] Brooks, R. A. (1990). Elephants don't play chess. In Maes, P., editor, *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, pages 3–15. MIT Press, Cambridge, MA.
- [8] Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1–3):139–159.

- [9] Bryson, J. J. (2000). Making modularity work: Combining memory systems and intelligent processes in a dialog agent. In Sloman, A., editor, AISB'00 Symposium on Designing a Functioning Mind, pages 21–30.
- [10] Bryson, J. J. (2001). Intelligence by Design: Principles of Modularity and Coordination for Engineering Complex Adaptive Agents. PhD thesis, MIT, Department of EECS, Cambridge, MA. AI Technical Report 2001-003.
- [11] Bryson, J. J. (2008). Embodiment versus memetics. *Mind & Society*, 7(1):77–94.
- [12] Bryson, J. J. (2009). Representations underlying social learning and cultural evolution. *Interaction Studies*, 10(1):77–100.
- [13] Bryson, J. J. (2010). Cultural ratcheting results primarily from semantic compression. In Smith, A. D. M., Schouwstra, M., de Boer, B., and Smith, K., editors, *The Evolution of Language 8*, pages 50–57, Utrecht.
- [14] Bryson, J. J., Ando, Y., and Lehmann, H. (2007). Agent-based models as scientific methodology: A case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society, B — Biology*, 362(1485):1685–1698.
- [15] Carlson, N. R. (2000). *Physiology of Behavior*. Allyn and Bacon, Boston, seventh edition.
- [16] Carruthers, P. (2005). The case for massively modular models of mind. In Stainton, R., editor, *Contemporary Debates in Cognitive Science*, pages 205–225. Blackwell Publishing.
- [17] Castro, L., Medina, A., and Toro, M. A. (2004). Hominid cultural transmission and the evolution of language. *Biology and Philosophy*, 19(5):721–737.
- [18] Clark, A. E. and Kashima, Y. (2007). Stereotypes help people connect with others in the community: A situated functional analysis of the stereotype consistency bias in communication. *Journal of Personality and Social Psychology*, 93(6):1028–1039.
- [19] Darian-Smith, E. (2002). Beating the bounds: Law, identity and territory in the New Europe. In Greenhouse, C., Warren, K., and Merz, E., editors, *Ethnography in Unstable Places*, pages 249–275. Duke University Press, Raleigh, NC.
- [20] Dawkins, R. (1976). The Selfish Gene. Oxford University Press.
- [21] Dennett, D. C. (1995). Intuition pumps. In Brockman, J., editor, *The Third Culture*, pages 181–197. Simon & Schuster, New York.
- [22] Dornhaus, A. and Franks, N. R. (2008). Individual and collective cognition in ants and other insects (Hymenoptera: Formicidae). *Myrmecological News*, 11:215–226.
- [23] Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, 20(2–3):193–203.
- [24] Fodor, J. A. (1983). *The Modularity of Mind*. Bradford Books. MIT Press, Cambridge, MA.
- [25] Franks, N. R. and Richardson, T. (2006). Teaching in tandem-running ants. *Nature*, 439(7073):153.
- [26] Harnad, S. (1987). Categorial perception: A critical overview. In Harnad, S., editor, *Categorial perception: The groundwork of perception*. Cambridge University Press.
- [27] Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. Complex Systems, 1:495–502.
- [28] Horswill, I. D. (1993). Specialization of Perceptual Processes. PhD thesis, MIT, Department of EECS, Cambridge, MA.

- [29] Huber, L., Range, F., Voelkl, B., Szucsich, A., Virányi, Z., and Miklosi, A. (2009). The evolution of imitation: What do the capacities of non-human animals tell us about the mechanisms of imitation? *Philosophical Transactions of the Royal Society*, *B* — *Biology*, 364(1485):2299–2309.
- [30] Kenward, B., Rutz, C., Weir, A. A. S., and Kacelnik, A. (2006). Development of tool use in New Caledonian crows: Inherited action patterns and social influences. *Animal Behaviour*, 72(6):1329–1343.
- [31] LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon and Schuster, New York.
- [32] O'Brien, M. J. and Shennan, S. J. (2009). Innovation in cultural systems: Contributions from evolutionary anthropology. MIT Press, Cambridge, MA.
- [33] Perry, S. and Manson, J. H. (2003). Traditions in monkeys. *Evolutionary Anthropology*, 12:71–81.
- [34] Rao, R. P. N. (1999). An optimal estimation approach to visual perception and learning. *Vision Research*, 39(11):1963–1989.
- [35] Richerson, P. J. and Boyd, R. (2005). Not By Genes Alone: How Culture Transformed Human Evolution. University Of Chicago Press.
- [36] Samuels, R. (1998). Evolutionary psychology and the massive modularity hypothesis. *The British Journal for the Philosophy of Science*, 49(4):575–602.
- [37] Sperber, D. and Hirschfeld, L. (2006). Culture and modularity. In Carruthers, P., Laurence, S., and Stich, S., editors, *The Innate Mind: Culture and Cognition*, volume 2, pages 149–164. Oxford University Press.
- [38] Sperber, D. and Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, 8(1):40–46.
- [39] Steadman, L. and Palmer, C. (1997). Myths as instructions from ancestors: The example of Oedipus. *Zygon*, 32(3):341–350.
- [40] Steels, L. and Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences*, 28(4):469–489.
- [41] Steels, L. and Kaplan, F. (1999). Bootstrapping grounded word semantics. In Briscoe, T., editor, *Linguistic evolution through language acquisition: formal and computational models*. Cambridge University Press.
- [42] Tomasello, M. (1999). The Cultural Origins of Human Cognition. Harvard University Press, Cambridge, MA.
- [43] van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S., and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*, 299(5603):102–105.
- [44] Waxman, S. R. and Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, 29(3):257–302.
- [45] Whitehouse, H. (2002). Modes of religiosity: Towards a cognitive explanation of the sociopolitical dynamics of religion. *Method & Theory in the Study of Religion*, 14(3–4):293–315.
- [46] Whiten, A., Goodall, J., McGew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., and Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399:682–685.

- [47] Whiten, A. and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. Advances in the Study of Behaviour, 21:239–83.
- [48] Whiten, A. and van Schaik, C. P. (2007). The evolution of animal 'cultures' and social intelligence. *Philosophical Transactions of the Royal Society*, B — Biology, 362(1480):603–620.
- [49] Wilensky, U. (2011). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/.
- [50] Wilkinson, A., Kuenstner, K., Mueller, J., and Huber, L. (2010). Social learning in a non-social reptile (geochelone carbonaria). Biology Letters, 6(5):614–616.
- [51] Wimsatt, W. C. (2009). Memetics does not provide a useful way of understanding cultural evolution. In Ayala, F. J. and Arp, R., editors, *Contemporary Debates in Philosophy of Biology*, pages 273–292. Wiley-Blackwell, Oxford, UK.
- [52] Wolpert, D. H. (1996a). The existence of *a priori* distinctions between learning algorithms. *Neural Computation*, 8(7):1391–1420.
- [53] Wolpert, D. H. (1996b). The lack of *a priori* distinctions between learning algorithms. *Neural Computation*, 8(7):1341–1390.
- [54] Wood, M. A. (2008). *An Agent-Independent Task Learning Framework*. PhD thesis, Department of Computer Science, University of Bath, United Kingdom.