Human Culture and Cognition

By

Matt Gers

A Thesis
Submitted to the Victoria University of Wellington
in Fulfillment of the Requirements for the Degree of
Master of Arts
in Philosophy

School of History, Philosophy, Politics and International Relations

Victoria University of Wellington
2008

Abstract

Human behaviour is largely influenced by culture. Culture evolves cumulatively over time. The origins of culture in our lineage necessitated the evolution of psychological biases so humans could tractably navigate the emerging information environment. I examine the nature of these biases and conclude that they are unlikely to be genetically coded to any significant degree. This is because of the flexibility such biases needed to possess in the face of fluid cultural environments and because of the developmental mechanisms of the brain. I further outline three possible views on what the nature of the information these biases act upon might be. First there is the view that cultural information is constructed and held in individual minds but does not flow in any meaningful replicative fashion between minds. Second is the view that culture is information distributed in a population and cultural evolution is the temporal change of this population-level information as a result of low fidelity individual copying events. Finally, I argue that meme theory, which asserts that culture is usefully seen as bits of information that replicate in transmission, is a fruitful model of cultural evolution.

1

Contents

Chapter One: Introduction to Culture

1.1
Why do people?

On the first page of his 1976 book *The Selfish Gene*, Richard Dawkins asks the question 'why are people?' The question my thesis addresses is the intimately related question 'why do people?' What follows is an analysis of some aspects of the cultural causes of human behaviour and the evolutionary forces acting on those proximate causes.

Humans appear to be purposeful rational agents, at least some of the time. Human behaviour is often a neat fit with the prevailing problems faced by human agents. For example Inuit natives living in harsh conditions dress in furs, fish in ice holes, and practice infanticide when resources are scarce. Observing such clearly adapted behaviour one might be tempted to ask, 'why are we so clever?' However, at other times human behaviour is counter-intuitive, harms the agent, or is plainly maladaptive. The problems of suicide, some religious rituals, polluting cultural practices, and even academic pursuit (as suggested by Richerson and Boyd 2005) may fall into this category. I believe the question, 'why are we so dumb?' is equally interesting and I plan to address both these questions in what follows.

I will argue that understanding human behaviour depends to a large degree on understanding the transmission dynamics of socially acquired information. This information forms the basis of much of our behaviour. I argue that this information is transmitted as mental representations of patterns among humans in, and between, social groups.

Patterns exist in our social worlds and the way we represent and respond to these patterns is fitness determining for human agents. But, these

representations are evolutionary replicators in their own right. They survive not only because agents that tend to acquire them survive to reproduce more frequently than agents who don't acquire certain patterns, but because pattern representations underlying social learning have fitness values of their own. Their environment is one of human psychology among other features. Furthermore, we possess psychological mechanisms, which bias the uptake of information from other humans and from artefacts. These biases constitute a strong force that dictates to a large degree what we learn from each other and from the pool of human culture.

I will also argue that these biases are context-specific and likely to vary from cultural group to cultural group. By arguing that human minds are not likely to be innately specified to any significant degree, I deny that genes code for domain general psychological biases, arguing instead that cultural transmission and the construction of niches cause certain biases to develop in certain conditions. When we appreciate what the function of the mind has been over evolutionary time, we see that the scope for innately specified biases is small indeed.

If these arguments in favour of replicating cultural entities (memes), and arguments against the predetermination of psychological biases are approximately correct, this will open the door to empirical work allowing us to predict which ideas, practices or artefacts will persist in human groups and also to overcome maladaptive biases through constructive developmental processes.

Culture is Information

In order to act in a purposeful manner an agent depends on information. Behaviour without an information basis is nothing more than random. Even the movements an amoeba makes as it engulfs a food particle are based on information. There is genetic information in its DNA; various cell-surface

receptor proteins are coded for by genes. These molecules react to objects in the environment and the information they transduce causes cytoskeleton rearrangements and purposeful phagocytosis by the amoeba. Information is critical to useful behaviour. Indeed, it is proposed that shifts to more complex organisms or behaviour over evolutionary history followed changes in the way information is acquired, stored, or transmitted (Maynard Smith and Szathmary 1995).

There are at least three information storage systems relevant to humans (Laland and Odling-Smee 2000). This is a controversial claim because the information content of DNA, for example, is contested (e.g. Moss 2003). But for the purposes of my argument I am assuming that the following three information systems exist. First, we have genetic information like all organisms. This information is the culmination of three billion years of natural selection on our planet. But there also exist ontogenetic information mechanisms, examples of which are the immune system and other learning mechanisms. The immune system produces an abundance of different cell-surface antibodies and, in response to extant pathogens, remembers which immune cell populations to sustain. Through other forms of solo learning an agent may also acquire and store information about his environment, for example as neural information. The third information mechanism is that of *culture*.

The definition of culture is tricky. In some senses culture refers to the symbols that social groups use in order to self-identify. Sperber and Claidiere (2006) argue that culture is better viewed as a property than a thing. In their view some things are *cultural* and some things aren't. The definition I will make use of here, however, is that culture is information acquired at least in part by social learning. This is broadly construed. Some authors make a further distinction between the cultural transmission of behaviour and that of symbolic culture (Jablonka and Lamb 2005). While this may be relevant when

considering the emergence, persistence, or relevance of particular cultural units, the conclusions of the present argument are not altered by considering all cultural traits as instances of the same general type, information learned by social means.

These three information systems, genetic, ontogenetic, and cultural, all contribute to human behaviour. We exhibit some behaviours, such as the primitive blink reflex, which are clearly genetically encoded. Genetic information generally takes a comparatively long time to establish itself by natural selection. Ontogenetic learning and immune systems can direct behaviour in a more plastic fashion than genes. The response of ontogenetic systems to environmental problems is generally fast. Culture has the potential to acquire information useful to the organism at an even faster rate again, as it is possible to bypass (or at least accelerate) the individual learning process and merely copy what someone else does. Culture permits flexibility of behaviour over and above an organism's genetic restrictions and asocial learning mechanisms.

In this thesis I focus upon the third of these information systems. What follows is an evolutionary explanation of culture and some aspects of human cultural evolution. Ultimately I attempt to outline part of an answer to the question posed above: 'Why do people?'

1.2
Précis of the argument

I believe that what human culture needs is *an explanation*. I am going to agree with Peter Richerson and Robert Boyd (2005, pg 237) who have said that, 'nothing about culture makes sense except in light of evolution'. We are evolved organisms, our information storage and acquisition devices are the result of natural selection, and it seems likely that evolutionary processes are at

work shaping culture too. I will be presenting and defending an evolutionary approach to culture. I will conclude that human culture is one aspect of an extended Darwinian paradigm (see appendix, figure 1).

The structure of the thesis is as follows. In the remainder of the first chapter I will introduce the essential concepts of cultural accumulation, the adaptive nature of culture, the importance of observational learning, and the notions of information cost and the information environment. Some brief discussion of these concepts is essential before moving into the argument proper. The main points in chapter one will be as follows:

> 1.3. Human culture accumulates.
>
> 1.4. Human culture is an adaptation and is adaptive.
>
> 1.5. Human culture relies to a large degree on observational learning.
>
> 1.6. Acquiring cultural information has a cost.

I will then argue for the following claims in the remainder of the thesis:

1. The complex nature of the information environment necessitates the evolution of psychological biases. These biases let agents tractably navigate this environment.

2. The upshot of debates in evolutionary psychology and neuroscience is that we cannot assume that these psychological biases are genetic adaptations. The necessary biases are likely to be partially or fully environmentally coded and vary across groups.

3. The vehicles for the information that these biases deal with are largely discrete replicating entities, or memes.

I argue for (1) in chapter two. I begin with a simple conceptual model of the origins of human culture, which provides insights into the changing nature of the information environment and the array of methods available for dealing

with this complexity. I then outline the mathematical models of Richerson and Boyd. Such models suggest that conformity and success biases are important aspects of human cultural evolution. But I note that these models have their flaws and it is unclear exactly what role these modelled biases play in human cultural evolution. I also note that the cumulative nature of culture is important, as is the structure of the cultural group itself. I discuss some likely transitions in the evolution of culture and conclude that the transmission and accumulation of cultural information over time necessitates cognitive biases of *some sort* in order for agents to solve the problem of adaptive cultural choice.

The likely nature of the cognitive biases that act on cultural transmission is explored in chapter three. It is possible that the biases are universal genetic adaptations, but it is also possible that they result from constructed cultural niches. I examine arguments from evolutionary psychology and neuroscience in order to try and constrain the possible nature of the biases. I argue that neurobiological and evolutionary constraints suggest that the massively modular adapted mind posited by evolutionary psychology is the wrong way to look at human cognition. I provide an overview of an alternative architecture, which solves the paradox generated when we realize that unbiased general purpose learners are very slow and that the human brain looks like a general purpose device. I then argue that in the face of this evidence the psychological biases that funnel social information must be socially learned or result from interaction with our constructed niche. Finally I provide evidence that this is indeed the case, therefore supporting (2) above.

Having described the psychological adaptations for culture we are left with the problem of what culture itself actually is. Chapter four provides a summary of three competing approaches to conceptualizing human culture. First there is Dan Sperber's 'epidemiology of representations'. This approach views culture as information that is transformed in every transmission event by rich evolved psychological attractors. I will argue that Sperber's approach is *prima facie*

wrong. Second I outline the population dynamic approach of Peter Richerson and Robert Boyd. Their view is that culture is copied with low fidelity in individual transmission events, but that the population distribution of information is inherited in stable fashion from generation to generation thanks to averaging and conformity mechanisms. Preservative replication manifests at the population level. The main competing theory to the Richerson and Boyd approach is memetic theory, which treats cultural traits as discrete replicating entities copied with high fidelity. I argue that there is a strong positive case for memetics that demands our attention.

In chapters five and six I defend memetics. I will present and diffuse the important objections to memetic theories. I will also explain what memes are and how they are replicated. This will involve discussion of Dan Dennett's theory of the nature of patterns as I argue that memes are mental representations of patterns existing in the world, therefore, supporting (3) above.

The result will be that, although Richerson and Boyd's models demonstrate the feasibility of evolutionary theories of culture, and suggest the necessity of psychological biases, the correct evolutionary approach is the memetic one and the biases are filters on memes.

In chapter seven, I present an overall synthetic theory of human culture that sees transmission as high fidelity. Variations in cultural traits are retained because of features of memes themselves and their appealing relation to the psychology of particular human populations, and also because of ecological factors and the effects of competing and complimentary memes. This information flow is funneled by often-adaptive cognitive biases that result from human social environments and niche construction.

I follow the discussion with a summary and possibilities for further research.

1.3

Cumulative culture

Human culture accumulates. Today we are completely immersed in the products of human culture yet 250,000 years ago there was considerably less. We see in many places evidence of the serial cumulative changing of artefacts. For example, human motor vehicle technology began with wheel technologies, then there were horse-drawn carriages, steam powered cars in the late nineteenth century, the internal combustion engine, streamlining, electric technologies, and safety improvements like air-bags. The lineage of car technology has become more complex, it has accumulated over time. The same is true of ideas, behaviours, languages, games, and institutions. Legal systems evolve incrementally as precedents are challenged and amendments made. Games like chess or rugby change and accumulate rules incrementally over time, and practices like the baking of bread have increased in complexity from simple flour and water flat breads to complicated loaf recipes involving yeast, sugar, butter and milk.

If an agent learns a behaviour socially and then this behaviour is altered by the agent so as to improve it, or increase its complexity, then culture can be said to have accumulated. It should be clear that cumulative culture presents the possibility of cultural ratcheting. If an agent improves upon a cultural practice then other agents can copy the *new* behaviour and possibly modify it further. If modifications are improvements then the practice improves incrementally. A lot of human culture exhibits this pattern. Early weapons for example were just sticks, later the tips were hardened in fires, later still stone heads attached. The spear was a product of accumulating culture.

Tomasello (1999) has argued that *imitation* is crucial to this process. It is possible that the explosion of cultural diversity seen in the last 250,000 years may have been a consequence of our acquisition of the ability to imitate. If

imitation is important then we must be able to tell a story linking the emergence of imitation to the current complex world of culture. I attempt such a story in chapter two.

The lineage patterns of many human cultural objects look exactly like phylogenetic patterns in biological evolution. This seems to be the natural result of cultural ratcheting. We are left wondering, however, how do we *focus* any imitative skills we might possess? What is the target of our social learning? Also, culture clearly changes and accumulates, but are these changes adaptive and are they adaptations?

1.4
Adaptive culture

Much human culture is clearly adaptive. The Inuits' practices help them to survive in extreme climates and elsewhere the replanting of forests will hopefully reduce starvation due to top soil erosion. But the adaptationist debate that suffuses the theory of genetic evolution has its cultural counterpart. It can be argued that much culture is maladaptive. We see the celibate practices of the Shakers, and African female genital mutilation rituals as examples. However, as a general rule Sterelny (2006b) argues that, 'the adaptive fit between human cultures and their environments is persuasive evidence that some form of evolutionary mechanism has been important in driving human cultural change'. An example of this fit would be the Balinese water temple system where the religious rituals harbour and disseminate information necessary for optimizing rice growth in a tropical climate. Also, as Sterelny (2007) points out, behaviours that were at one time adaptive can dead-end in 'fitness traps' with a less-fit valley on the adaptive landscape separating the behaviour from a fitter behaviour. A fitness trap is a situation where an individually advantageous strategy invades the population but when it becomes universal it reduces the absolute fitness of every individual. An example of this

might be the playing of a violent sport. When a few individuals do it their desirability as mates may rise. The practice is adaptive because they are advertising survival skills. But once everyone in the population is taking part then everyone's fitness drops. This is because of the continuing risk of injuries with no increase in relative desirability. Therefore, the existence of maladaptive culture does not mean that most features of culture are not adaptations. Indeed, the conceptual and mathematical models that I will discuss in the next chapter demonstrate that culture is indeed adaptive even though particular cultural practices may at times be maladaptive.

Culture will be adaptive when learning is expensive and difficult, and environments are variable, but not *too* variable (Richerson and Boyd 2005). If environments are stable then fit behaviour has time to become genetically entrenched. If environments vary too much then it pays to learn on your own because of the very real risk that anyone else's strategy may already be out of date. Some models of the advantages of cultural transmission in various circumstances, therefore, show that in stable environments an agent ought to learn from their parents and in unstable environments they ought to learn on their own. Of course intermediate possibilities exist. For example, a ten year old might do well to copy a teenager who has recently grown up in the current environment rather than copy his out of date parents or spend a lot of effort learning on his own by trial and error. We need to distinguish within and between-generation copying. If the members of generation 2 are copying their parents then this is vertical transmission of culture. If they are copying their peers then this is horizontal, or within-generation, transmission. Oblique transmission is possible too. This is where an agent copies from non-parental members of the previous generation. But how does the agent choose which strategy to employ? Early models of cultural learning by imitation (e.g. Rogers 1995) miss these crucial aspects of culture – how to choose who (or what) to imitate, and the costs of determining which behaviours are successful. Being picky about who to learn from can confer the advantages of both imitation *and*

learning (Richerson and Boyd 2005). It is learning in this choosy or biased fashion, along with imitation, which are the adaptations from the point of view of human agents. I will examine the existence of such psychological biases in chapter two. In chapter three I will examine their nature, and in chapters four through six, I will examine what it is that these biases are actually acting upon, what the cultural 'bits' are that are actually funnelled by these biases.

1.5
Three ways of learning by observation

I will now explain the importance of observational learning in cultural processes and detail three possible ways that information can be learned by observation.

Many animals exhibit practices that might be called animal culture (Avital and Jablonka 2000, Laland forthcoming). However, even intelligent, social apes don't have anything remotely like the human capacity for culture. Therefore, as Richerson and Boyd rightly point out (2005), more than a big brain and social organisation are required for complex culture. To take an analogy, red blood cells cannot reproduce themselves because they lack a nucleus and the apparatus of DNA transcription. For the spread of a cell-line, more is required than a membrane system, organelles and the various other machinery of cell function. In the case of culture it seems clear that there must be some special purpose mechanism (or indeed mechanisms) other than a large brain and social organisation that enable behavioural practices as complex as those of human beings to spread culturally. Also, we cannot assume that it is merely some psychological device that is missing in the great apes. It may be a combination of psychological factors, social structure, group size or density, even ecological factors that have led to the take off in human culture. What is clear, however, is that the capacity for observational learning is of significant importance.

Seldom do we see cumulative culture in nature. Possible exceptions are song dialects in some birds and some chimpanzee behaviours (Boyd and Richerson 1996). For cultural practices to accumulate, that is, for the practice of agent B to build upon the practice of agent A, then agent B must actually observe and copy the practice of agent A, or at least get the idea from agent A and then fill in the details by his own trial and error. In such hybrid learning, if there is an optimal method, pure imitation, and trial and error according to a learned template, ought to converge on the same behaviour. If agent B only learns on his own, with no input from agent A, then this cannot lead to culture in any form.

It is debatable whether such learning mechanisms as stimulus enhancement and local enhancement suffice for culture. In these cases the environment of the learning agent is such that he is more likely to learn certain behaviours than others. For example the practice of termite fishing by chimpanzees may be learned not by observing older chimps, but because the infants spend a lot of time around sticks and termite mounds while their mothers are fishing for termites. Whether these processes count as culture or not, it appears that they cannot sustain the *accumulation* of cultural practices. Richerson and Boyd (2005) claim that *observational learning* seems to be essential to cumulative cultural evolution. As noted above, the learning must be *from* agent A.

It is possible that accumulating modifications of the environment could sustain some cumulative social learning. For example the accumulating discoveries, over generations, of how to tame fire, keep a wood supply, build a protective hearth, cook food, and draw with charcoal could be sustained by the effects these discoveries have on the immediate environment of the cave. Such niche construction means that the learning environment of subsequent generations is biased in favour of learning these behaviours that no one individual could have

learned on their own. This may suffice for some accumulation of culture, but it is only a part of the explanation for human culture.

Novel practices can indeed change the trial and error environment (Avital and Jablonka 2000), so in the above example the fact that young chimps spend an inordinate amount of time around termite mounds may lead them to learn how to fish for termites more quickly than the previous generation and therefore the younger chimps may be more likely to *improve* upon the practice. But we might still claim that observational learning is necessary for *rapid* or *complex* accumulation of cultural practices (as seen in humans). So what is observational learning? I will now outline three methods that may sustain cumulative culture: imitation, emulation and teaching.

Imitation

Imitation is a tricky term (Laland forthcoming). It is difficult to isolate the many complex mechanisms at play in the folk psychological concept of imitation (Chaminade forthcoming). Imitation may, however, be defined narrowly as *reproducing through observation the specific motor pattern of another individual*. There is no doubt that humans seem particularly good at imitative learning and human infants persist in imitative techniques even when chimps abandon them in favour of contextually successful trial and error learning. We are good at imitation and biased to do it (Richerson and Boyd 2005). But to what degree this constitutes our cultural abilities is less clear. Also, there exists a particular sort of imitation, which we might term *program-level imitation*. In these cases it is the general organization of the behaviour, not the exact motor pattern, which is copied. Here's an example. If your mother is standing on a ladder reaching up into the orange tree you might *imitate* her by finding another ladder just the same and reaching upwards in a similar fashion with the same hand. This would be narrow imitation. Or, you might infer that she is trying to reach the oranges and that the program of

behaviour is a sequence of first obtaining a stepping device, then reaching up with whatever hand is nearest the orange. This would be program-level imitation.

It ought to be noted here that Richerson and Boyd frequently refer to 'imitation' in their cultural models, but it is not always obvious what they mean precisely. It often seems to be the case that they are discussing *preservative copying processes* more broadly and not always narrow imitation as defined here. This must be borne in mind when we come to discuss their theories.

Emulation

Rather than copying the exact motor pattern or general organization of a behaviour (tasks which do not necessitate understanding the goal of the movements) an agent may infer the intent of the target and construct a motor pattern that achieves the same outcome. This process may be called *goal emulation*. Some authors (e.g. Laland forthcoming) believe that emulation is far more common in humans than narrow imitation. There *is* evidence for narrow imitation in humans. For example, we imitate the body language of those we communicate with to an impressive degree. Also, human infants seem to be able to imitate facial expressions from between 12 and 21 days old (Meltzoff and Moore 1977). However, babies then lose this ability so it is questionable whether it has relevance for studies of cultural learning. It seems that more often we emulate with a goal in mind rather than merely copy motor patterns. I note also that emulation is a form of hybrid learning where the individual combines social and asocial techniques in order to acquire a novel behaviour. As pointed out above, such hybrid learning may result in high fidelity cultural transmission depending on the other constraints imposed upon behaviour (e.g. optimal strategies, available tools, background knowledge). Here's an example. If you were to merely *emulate* your mother's actions

above, you might climb the tree, rather than using a ladder and, therefore, achieve the same goal by quite different means. Similarly, you may attempt to copy Roger Federer's tennis serve. Unless you have a slow motion camera you cannot pick out the exact body movements that he makes, however, you can see that he tosses the ball in order to be able to hit it over the net. You may copy this program and intent and achieve the goal to a greater or lesser degree, a degree that may be constrained by your physiology, strength, co-ordination and so forth.

Teaching

Teaching is another tricky term (to again borrow Laland's phrase, forthcoming). He defines teaching as *costly modification of behaviour specifically designed to instruct a tutee*. In some sense teaching is a formalization of the process of editing a behaviour, in order to make it better fit the problem faced. The formalization involves instruction, and observation of the resulting behaviour, by a skilled practitioner.

So there are three ways in which we can acquire information by observation. Like Richerson and Boyd, however, I will use the single term imitation from now on to refer to these preservative observational learning processes generally.

1.6
Cost and the information landscape

I have just mentioned cost with respect to behaviour. This will become an important concept in chapter two where I examine the origins of human culture and the psychological biases that funnel cultural transmission. So I will take some time in this section to explain the concepts of the cost of information and the information environment.

Ordinarily we talk about something's cost as being that which one must forgo in order to obtain it. We can talk of the cost of acquiring information as those resources, for example time, material goods, potential alternatives forgone, in order to obtain the piece of information. Also we need to include any 'running costs' incurred by the agent while pursuing the information. These may include metabolic expense or other measures of effort. A 'costly behaviour', therefore, or one which is *expensive* to obtain, is one for which an agent forgoes a lot of valuable resources to acquire. There are also risk costs. The more chance there is of a negative outcome the more costly an acquisition event becomes. Generally evolution favours low-cost methods because this maximizes the agent's reproductive chances and fitness.

We can also talk of the *information environment* of human agents. This is the sum total and distribution of all potentially copyable, or learnable behaviours, or semantic knowledge on offer in an agent's environment. We can metaphorically speak of the *information landscape*, and of an agent *navigating* such a landscape as they pay selective attention to some information and not to other information. Some behaviours may increase the agent's fitness to navigate the information environment by causing the agent to tend to acquire useful rather than useless information. We may also talk of the information *flow* from agent to agent or across a population. Such flows are merely patterns of transmission of information from one mind to others. Dan Dennett (1998) introduces the idea of an *information gradient*. This is where the information held by agents differs. This may enable complimentary or exploitatory interactions. But in either case communication will be an essential element of human interactions. Of note the information gradient in human environments has steepened over time.

Finally, we can abstractly represent all possible behaviours on offer as a region of state-space each point in the state-space rising to a peak that represents

fitness. The dimensions other than height represent behaviours in the prevailing environment. The taller the peak the better the behaviour for the agent's survival. The evolutionary process can 'climb' peaks by natural selection but cannot jump peak to peak across valleys because behaviours in between are less fit and will be selected against.

1.7

State of play

I have shown that culture accumulates, that human culture is adaptive, and that rapidly accumulating culture relies on observational learning mechanisms (hereafter referred to as imitation). The origin and nature of these features of culture need an explanation. I offer such an explanation in the remainder of this thesis.

I have also explained how social information can vary in its cost and how culture results in an environment of information that must be navigated by the agent. These features of culture will act as constraints on the dynamic of cultural transmission and I explore these constraints in the chapters that follow.

Chapter 2: The Information Environment and the Evolution of Psychological Biases

Overview

In the first chapter I explained that the emergence of social learning or imitation permitted the cultural flow of information among humans. This development in our lineage has had far reaching implications. Our minds and social structures have responded to the emergence of culture with an array of mechanisms for producing and retaining adaptive benefits from culture. In this chapter I will first demonstrate with a conceptual model how the emergence of imitation led to the persistence of adaptive culture. I will then show that further insights about culture can be gleaned from mathematical models, which permit us to see the results of iterated processes. Boyd and Richerson believe that their mathematical models show that a conformity bias and prestige processes can together result in the retention of complex adaptive culture in a population. I will show why there is good reason to be very cautious when interpreting their models, although I will agree with their general assessment of the need for psychological biases. I will argue that biases for channeling culture are an essential development for a population that begins to accumulate a complex culture as a result of imitative capabilities, but that there is more to the biases than Boyd and Richerson claim. Their analysis is not evidence enough to account for the features of culture that we see. I will argue that both conformity and prestige biases can be maladaptive at times, and that there is no reason to believe that they are important in all cultures in all times. I will also show that structural features of populations and epistemic technologies are important in retaining complex adaptive culture. Imitation enables culture and culture triggers the emergence of biases and structural changes in populations. But biases are unlikely to be universal and must be developmentally plastic.

2.1

Introduction to early culture

Hopefully I have established that information is transmitted between human individuals. I now turn to the implications of such a process for the evolution of human psychological mechanisms and the accumulation of culture. My central point in this chapter will be that there is a dynamic flow of information between humans and that heuristics, biases and social structures influence this flow in adaptive ways. Societies and minds *respond* to the increasing availability of information.

First I will outline a conceptual model of information transmission between early humans. I will then return to the idea of cumulative culture, and illustrate how the natural ecology of humans, the information environment, becomes very complex. I argue that humans *need* a series of psychological heuristics in order to tractably navigate this environment. In the next chapter I will discuss the nature of these biases. In this chapter I merely argue that the evolution of these heuristics is an entailment of our life in the cognitive niche and the pervasive algorithm of natural selection.

Modern human culture did not suddenly emerge from some pre-cultural hominid ancestor (Plotkin 1996). The emergence of behaviourally modern humans is distributed through space and time (Henshilwood and Marean 2003). These authors describe the possibility that the pre-historic appearance of technological complexity in Europe and in Africa may have occurred in response to processes unique to each context. There may well have been a point origin for a culturally capable individual, but if so then it seems that his or her lineage had spread before complex culture began to arise in different locations. It would follow then that cultural capability were an exaptation, which was put to cultural use when conditions were ripe for it. This is not unreasonable as many of the suggested prerequisites for culture have obvious

alternative applications such as theory of mind or joint attention. The different locations in which culture arose may show different patterns of cultural development and accumulation because, even if culturally capable, agents may not develop complex culture unless the dynamics of group size and population density are favourable. Larger groups allow for specialization and redundancy, both of which enhance cultural accumulation.

Culture seems to have accumulated in several locations over a relatively short evolutionary time span. I will argue that this may have led to the emergence of an array of *similar* cognitive biases in different locations. This will be a theory of the convergent evolution of biases. Biases will look similar and function in the same way like sharks' fins and whales' fins. But the question of whether these biases arose because of biological selection or were due to cultural selection will remain. I address this further issue in chapter three.

The discussion in this chapter will address these issues:

1. A conceptual model of the adaptive nature of culture
2. Mathematical models of cultural evolution
3. The accumulation of culture
4. The increasing complexity of the information environment
5. The response of minds to this complexity:
    a. Non-content cues
        i. Conformity and Gullibility
        ii. Prestige processes
    b. Content cues
6. Group structure and crowd cognition
7. Transitions in the vectors determining cultural evolution

In 1933 the Kellog's decided to raise their infant child with a chimpanzee. They expected the chimpanzee, renowned for its 'imitative' behaviour, to learn many human-child-like behaviours by imitation. As it turned out the chimpanzee learned very few child-like behaviours whilst the child imitated and learned an array of chimpanzee behaviours such as knuckle walking. It appears that humans have a remarkable capacity for imitation which is absent in even our closest great ape relatives. I will assume that the ability to imitate other agents (in the Richerson and Boyd sense of preservative observational learning) became massively amplified at some point in the human lineage. This is a necessary precondition for the copying of cultural traits. It is unclear when this occurred but some evidence suggests it may have been relatively recently that humans evolved the ability to copy with high fidelity. The evidence for this is multifactorial. Not only do the great apes lack *significant* skill in imitation, but human stone tool technologies were static for long periods from around 2.5 million to 250,000 years ago. This suggests that improvements were not retained. Tomasello (1999) emphasizes the importance of imitation in permitting cumulative cultural improvement, but true imitation is cognitively sophisticated (Sterelny 2003a, Distin 2005). It is only once this ability is in place that we see cultural processes really get off the ground. What concerns us next are the implications of this ability for the human lineage. Imitative copying clearly permits social learning, and social learning is adaptive and persists. Why might this be so?

2.2
A conceptual model of early culture

I'll refer to the following as the *Simple Model of Cultural Origins*. Let P be a population of non-cultural agents of individuals p1, p2, p3... etc. These agents have a repertoire of genetically programmed behaviour and can also learn by

individual trial and error. This means they have both genetic and ontogenetic information systems at their disposal. A mutant (p-m) arises in P. This model holds if the mutant is a novel mutant or if they are a migrant from another population P* in which a novel mutant originated. The mutant possesses a rudimentary ability to imitate. He can learn by observational methods and so can be classed as a social learner. The ability to copy the behaviour of others permits a particular sort of behaviourally relevant information to be transmitted between individuals. Note that in this model I am greatly simplifying the situation. Parallel processes such as niche construction may enable stable social learning without imitation. Imitation here merely means preservative copying processes generally.

I assume that for individual agents a strategy of social learning is fitter than asocial learning since it reduces the costs involved in acquiring adaptive information. I assume that this is the case even if copies are partial (giving a 'leg-up' in the learning process) or a hybrid learning technique is employed in which the social learner copies some aspects of behaviours and then works out the rest on their own. I am assuming that what is copied is copied with sufficient fidelity that the function of the behaviour is usefully preserved.

I argue then that p-m will be at a fitness advantage relative to p1, p2, p3… and that he will have more offspring. Assuming that the offspring inherit the imitative capacity (and it is cheap relative to its benefits) then they, too, will be fitter than average and the trait will spread. Initially p-m and his offspring obtain adaptive or neutral information because extant behaviours are largely genetic adaptations or learned solutions to local problems. The new mutants will do well and the degree, quality, and reliance on imitation will increase (see figure 2.1).

Now the environment changes. At this point p-m numbers may be quite high. Their success now depends on the learning strategy employed. If they use a

copy-only strategy they risk obtaining behaviour adaptive in the old environment. In that case they will fare poorly and their numbers will decrease. This will be true unless they can learn a metastrategy of not always imitating. However, they will not be eradicated because when their numbers drop low enough they will be copying only from non-social learners who have obtained the newly adaptive behaviour by trial and error. The tendency to learn socially would seem to persist in boom-bust fashion in non-accumulative cultures (see figure 2.2).

a

b

Figure 2.1

In this diagram arrows within the boxes represent the flow of cultural information. The large box (P) represents the population of individuals p1, p2, p3… The mutant imitator is represented by the small p-m box. Here we see the success of the mutant imitator over time. Initially in box (a) we see the flow of cultural information (i) from the population P to the novel imitator p-m. After some time (box b) p-m's numbers have increased to a larger proportion of the population. Useful information (i) continues to flow to the members of p-m both from the original population and among the imitators themselves.

| a | b | c |

Figure 2.2

The environment changes and box (a) shows the resulting situation. Arrows still represent the flow of cultural information, but now we see that some information is maladaptive (m). Anything the members of p-m learn from each other will be of this kind. Some information (i) is still adaptive because members of population P outside the imitator box are still learning by trial and error. But as time passes the number of imitators decreases (box b) and the quality of information (i), now mostly obtained from individual learners, increases. The p-m population will grow again (box c) and persists in boom-bust fashion in a changing environment.

Let's stop assuming a copy-only strategy for p-m. If the social learner employs a mixed strategy of copy-and-learn then he can avoid the costs of the environment changing. He can copy behaviours still relevant in the new environment, and learn other new skills on his own. This mixed learner will need some strategy for deciding when to copy and when to learn individually. Part of the answer may be to copy only in unfamiliar situations. This would at least eliminate the drawbacks of copying a poor behaviour when he already knew what to do. Let's assume that this non-trivial problem is tractable. Overall, the optimal strategy will depend on the rate at which the environment changes. If it changes often then a copy-only strategy fails because there is a lot of out of date information around. If it doesn't change at all then a copy-only strategy is best because extant behaviour will be adaptive. In between cases favour a mixed learning strategy.

So assuming the environment is non-static then eventually individual learners give way to mixed learners and copiers. The population may now be called P-M. The best strategy now is again a copy-only strategy. Such freeloaders don't bother spending a lot of time and effort learning anything, if they get stuck they just imitate. The target of their imitation is either a behaviour learned by trial and error or it is adaptive behaviour copied by their mixed-learner model. Freeloaders will do well until the environment changes. Then they risk copying each other and obtaining out of date information. In this case they are also an irritating problem for the mixed learners who must learn to avoid copying freeloaders. Freeloaders also persist in boom-bust fashion in the population P-M because the quality of the information they obtain rises as their numbers fall (see figure 2.3).

The conceptual model I have just described, the Simple Model of Cultural Origins, is just that, simple for illustrative purposes. We have yet to address the issues of cumulative culture, information environment complexity, the evolution of heuristics and biases, and so forth. What is apparent, however, is that social learning is adaptive and persists given our initial assumptions about the costs of acquiring information. It also seems as though the pace of environmental change dictates the time-point composition of learning strategies in populations. We can make some preliminary conclusions.

1. Hybrid imitation-learning becomes the norm.
2. Good and bad information exists and agents must learn to discriminate.
3. Different strategies for navigating the information environment exist (e.g. freeloading, individual trial and error).
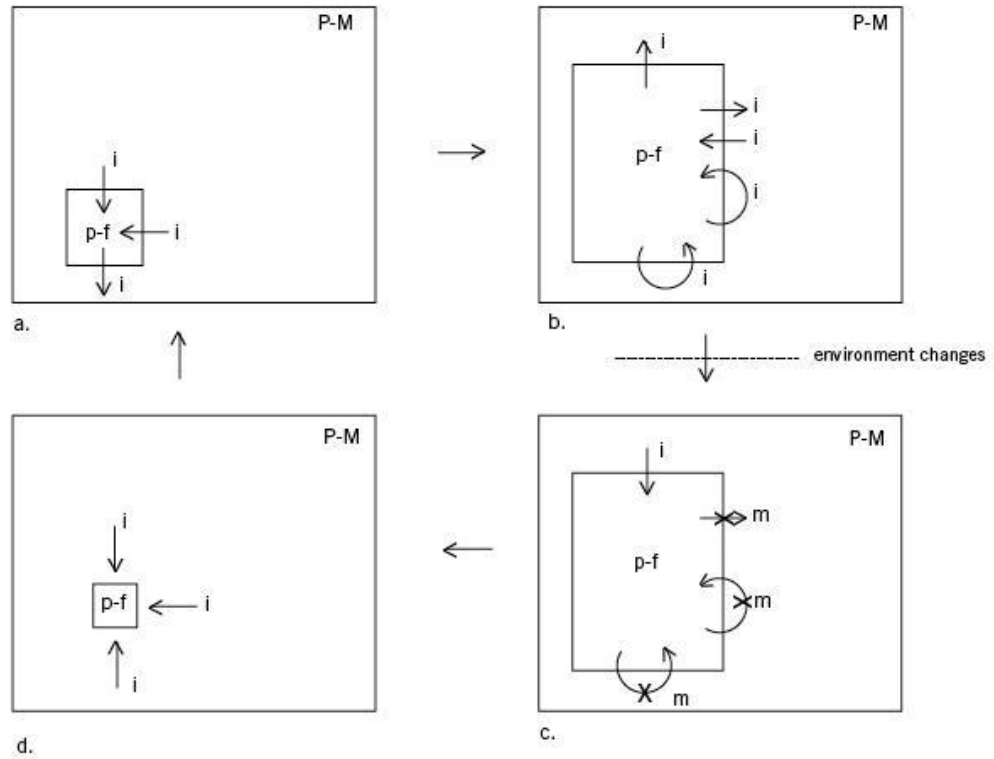
Figure 2.3

Note that once the entire population consists of social learners (P-M) then social information flows both into and out of a population of freeloaders (p-f) who initially fare well. This is a problem for the members of P-M if the environment changes (as occurs between box b and c). Members of P-M then risk obtaining maladaptive information (m) from what has become a large population of freeloaders. However, eventually freeloader numbers will decrease due to the poor information (m) they glean from each other. The cycle then continures.

Having argued that the emergence of imitation necessitates discriminative strategies, I now move on from conceptual models to mathematical models. Richerson and Boyd claim that mathematical models can generate hypotheses for what the discriminative strategies actually could be. Their models predict the evolution of conformity and success (or prestige) mechanisms. I will outline their mathematical approach, demonstrating how they reach the same insights as the conceptual model does and also produce new insights into cultural processes. However, I will find that we must be cautious in

interpreting their results. The theoretical processes of conformity and prestige-biased copying are much more complex in real world situations and are not necessarily as adaptive as Richerson and Boyd claim.

2.3

Mathematical modelling of cultural evolution

Shennan (2006, pg 2) criticizes exactly the sort of simple conceptual model above. He notes that:

> Words are all very well but only mathematical modelling and computer simulation can tell you what the outcome will really be when a certain process is repeated over and over again.

This is true; many of the most interesting findings in theories of cultural evolution are obtained through mathematical modelling. Richerson and Boyd claim that without mathematical models we are forced into using verbal arguments of questionable logical consistency (2005). It is time to discuss the merits of various methods of investigating early culture.

McElreath and Henrich (2007) formalize a model similar to the conceptual Simple Model of Cultural Origins above. By accounting for L, the change in frequency of individual learners in the population in the next generation, B, the payoff of the adaptive behaviour (they assume only one is adaptive), C, the cost of experimentation and mistakes (trial and error), and u, the probability each generation that the environment changes, they derive:

$$L = \frac{(B/C) - 1}{(1/u) - 1}$$

This demonstrates that the prevalence of individual learning in a population responds to the costs of learning and the instability of the environment. As B/C decreases, learning becomes costly. As u increases, individual learning at equilibrium increases in frequency. Indeed, as u tends toward 1 the denominator tends towards zero and L tends toward infinity. As we reasoned above, if the environment is changing, copying your parents doesn't work[1]. For environmental fluctuations overall, however, natural selection favours substantial amounts of social learning unless u is *very* large. This is not at all obvious from the conceptual model above. What if u is very small? As u tends to zero, L tends toward zero too. So if the environment never changes then social learning dominates[2]. In this mathematical model, unlike in the conceptual model, we can model the *frequency* of environmental change. We could also model the magnitude of environmental change, this would be a measure of relevant change (however, this model does not account for such degrees of change). Of course, there will always be *some* individual learning because the cost of learning will vary vastly depending on the problem. There will always be problems that can be solved by quick and simple direct perception. Agents will continue to do this.

If culture is accumulating then u does have the potential to be very large as the cultural environment can be changing rapidly. This will be relevant later when we examine the complexity of the information environment. We can predict that rapid cultural accumulation, particularly the non-linear accumulation of fitness enhancing technologies, will either drive more individual learning (rather than reliance on the information previous generations have to offer), or it will drive technologies that enhance relevant social learning channels. This may be by improving pedagogical methods and enhancing relevant horizontal

---

[1] This model assumes that social learning occurs with younger individuals learning from older ones (McElreath and Henrich 2007, pg 581). Such vertical transmission can, of course, be supplemented by horizontal learning from peers in real world situations. This may change the dynamics of the model.

[2] This is, of course, true for the mechanisms of social learning in general, however, Baldwinian processes would be expected to drive the proximate mechanisms guiding any *particular* behaviour into the genome if the environment is truly stable.

learning channels. It is possible that the internet is performing both these functions in the present environment. What is also apparent is that a group of social learners will do best to stabilize u and keep C as low as possible. If a *group* can absorb physical environment threats, stabilize cultural traditions, formalize the learning process and/or reduce the costs of individual learning, then they will rely less on costly individual learning and the fitness of the group ought to rise. At the level of the group we see a tension between technological advance, which is usually fitness enhancing but entails an increased cost of individual learning due to a rising u, and stabilizing forces on u to decrease the reliance on individual learning and therefore decrease the group costs of learning. Such a situation ought to drive selection for division of labour. If we bring the possibility of cultural group selection into the equation then the dynamics of the models have the potential to change significantly.

Some Cautionary Remarks Regarding Mathematical Models

I will now argue that there is a risk of over-reliance on the Richerson and Boyd-type models. This is often because there are important features of human cultural transmission that are omitted from or misconstrued in the formal models. It is true that Richerson and Boyd run iterated simulations, which can lead to counter-intuitive and surprising results. This teaches us not to rely *solely* on conceptual models. This is especially the case when it appears that opposing forces are acting at different levels of selection and also when individual activities feed up to population-level effects which in turn feed back as selective forces on individuals. This occurs with dual inheritance and niche construction models of human cultural evolution (Laland et al. 2000, Odling-Smee et al. 2003, McElreath and Henrich 2007; see also appendix diagram 1 below). Henrich and Boyd (2002, pg 109) note that:

> Our minds are not well equipped to understand population-level outcomes that arise from numerous minor interactions, weak cognitive biases, random errors, migration rates and micro-level decisions.

They intend this as a reason to rely on mathematical models when theorizing about cultural evolution. But on the other hand models run the risk of being engineered by those who lack the necessary theoretical grounding in a discipline and such simulations are also necessarily simplified. Simplification is not in itself a problem; much of science is just the simplification of complex interactions to straightforward models. Think of the firing of neurons modelled as an on or off process. But mathematical models must at least manipulate relevant variables and values that are at least in the vicinity of the actual systems they model. This is not always the case.

Here is an example that demonstrates that Richerson and Boyd may trust their models too much. This shows that we must be careful in evaluating the methods employed by mathematical modellers.

Henrich and Boyd (1998) describe a model of cultural evolutionary processes which indicates that a bias toward frequency-dependent copying (conformity) in social learning arises in all situations where a reliance on social learning itself evolves. This conclusion may be true for specific evolutionary systems, but it is difficult to assert from their model that this is the case for humans because the simulation models variables with values outside the range experienced by past human populations (see next paragraph for details). Also, the model is importantly incomplete and leaves out relevant variables and constraints. For example, below I argue that vertical transmission is important in cultural evolution and may corrupt simple models that privilege conformity biases.

The Henrich and Boyd model assumes a genetic system that has twenty alleles for varying degrees of the tendency to imitate, and twenty alleles for varying reliances on conformist social transmission. Each individual is therefore to some degree an imitator and to some degree a conformist learner. Initially

99.99% of the population are non-imitators and non-conformers. The remaining 0.01% are an even mix of all the other 399 possible genotypes. Here we strike the first problem with this model. For each genotype to be represented once then we need a total population of four million individuals. This is the population estimate for humans on earth as recently as 10,000 years ago (this is only in the order of 300 generations). It is not a realistic model of deep cultural origins.

Next Henrich and Boyd explain that in their model there is only one adaptive behaviour (realistically several variations may be adaptive to different degrees) and that the environment changes with probability ε between two alternative states, suddenly, between generations. Obviously this binary system of behaviours and environments with sudden reversals is a distortion of real environments which may change dramatically but in degrees over time. Finally, the simulation is run for up to 300,000 generations.

So what we have is a model of a population vastly larger than any pre-historical human population, and a simulation run over something in the order of six million years. The fact that humans have been capable of imitation for somewhat less than 250,000 years, and perhaps even less than that, is important. These numbers mean that a more realistic model would consist of fewer initial alleles, fewer initial individuals, and fewer generations of iteration. In some cases here 'fewer' means orders of magnitude! Henrich and Boyd accept that their model is a simplification, and they discuss relaxing some of their assumptions, but this relaxing often involves *increasing* the number of individuals. One could accuse Henrich and Boyd of generously assuming the conditions needed for conformist transmission to evolve in their model. They have established the answer that conformity evolves, but only given these arbitrary conditions. Changing the structure of the social group can have dramatic effects upon the take-off and persistence of adaptive culture. I discuss this under *crowd cognition and group structure* below.

What has been shown is that abstract genetic evolutionary systems do evolve conformity bias. But I question whether it happened like this in the human lineage. Mathematical models must be created and interpreted carefully. Simplifications are all very well when that which is excluded is less important than that which is modelled. The question is whether that which has been excluded in these models is important or not. I think it is.

I am about to argue that the accumulation of culture makes learning difficult and that complexity in the information environment does necessitate biases to help tractably navigate cultural worlds. Richerson and Boyd think that their models show us what these biases could be (conformity and success biases) and that these are (probably) genetic adaptations. But, in light of the present criticisms of their models we now need to carefully evaluate these biases in the context of human cultural transmission to see if this explanation is adequate. After examining the biases, and some structural features of cultural groups that contribute to the generation and retention of adaptive culture, I will conclude that the present explanation of culture is not adequate.

2.4
The accumulation of culture makes learning difficult (and also necessary)

Neither the conceptual model above nor the mathematical model addresses the issue of culture's accumulation. Once the imitator p-m possesses the ability to actually copy the behaviour of others then it becomes possible for him to save on initial learning costs and instead of paying a price to obtain adaptive behaviour he can pay a price to *improve* it. This means that in a population of imitator-learners there will be a lot of different behavioural variants on offer (again compare this to binary, or few-variable simulations). This also means that cultural groups will diverge in their practices as different innovations are created and spread in different groups.

As Richerson and Boyd note in *Not by Genes Alone* (2005), human culture is not merely social learning where one individual copies the exact behaviour of another, it is also an accumulation of novel innovations. As discussed in chapter one this accumulation leads to a highly adaptive culture (think again of the Inuit). Richerson and Boyd are skeptical, however, of the ability of imitation alone to preserve adaptations. They believe that our copying fidelity is too low. I disagree as I will discuss at length in chapter five. But I do agree with much of the theory that Richerson and Boyd favour to help explain the retention of adaptive culture. They agree that social learning is not sufficient *on its own* to explain cumulative cultural adaptation. They postulate, in addition, several heuristic mechanisms for extracting information from the population as a whole rather than from individual targets of imitation. These context-dependent biases will be discussed below.

So it can be seen that naïve agents have the following difficult problem. Given that there is an array of behaviour on offer, that there exist copy-only freeloaders with old information, that there is deceit as an evolutionary strategy, and that agent's need to retain adaptive behaviour while eliminating less-fit innovation, *which* behaviour do they copy in order to obtain fit information from a complex environment? This is the problem of adaptive individual choice.

2.5
The information environment is complex

All environments contain vast amounts of information, but some of this information is more relevant than the rest. For example storm clouds on the horizon are more relevant to whalers than to a seal. Advertising billboards are more relevant to me than to my pet dog. As the amount of relevant environmental information complexity increases (this may happen over evolutionary time as culture evolves or over developmental time as infants'

discriminatory powers increase) it becomes more and more difficult to evaluate the information because there is so much of it, and because its reliability and relevance is not stationary. Rules which take a few reliable proximate cues of good outcomes ought to be employed in order to save on evaluation costs. The evolved psychological biases, such as the conformity bias discussed by Henrich and Boyd (1998) are supposed to be exactly this kind of rule.

The natural environment of humans includes information conveyed by conspecifics. Bergstrom, Moehlmann and Boyer (2006) suggest that, therefore, we ought to find early-developed specific capacities for acquiring and transmitting information, as well as capacities for evaluating its reliability and for deceiving. Sperber realizes the importance of this point:

> The human reliance on communication is so great, the risks of deception and manipulation so ubiquitous, that it is reasonable to speculate that all cost-effective available modes of defense are likely to have evolved. (Sperber 2000b)

> Inconsistencies in perception-based information occur, but they are rare, and trying to eliminate them may not be worth the cost. For richly communicating animals like us, eliminating inconsistencies [in social information] may not only be worth the cost, but indeed, literally life-saving. (Sperber 2000b)

It seems that we should either possess robust innate mechanisms for solving this problem or that the primary goal of our parents, and genetically related kin with interests in our fitness, ought to be endowing us with such mechanisms. We need to become discriminative as to who to trust and sensitive to deceit.

Marie-Odile Taillard (2004) examines the inherent tension between human communicators and audiences. She notes that the function of communication is twofold, to obtain information and to persuade. The first goal is to learn more about the world without risk. The second is to effect changes in the world so that the state of affairs is nearer to the agent's desired state of affairs.

Interestingly, persuasive techniques will face diminishing returns as audiences learn to cope. It appears that in navigating the complex information environment through language, consistency of information is initially more important than its source. Especially in child studies (e.g. Harris 2002) the agent's focus is on whether acquired information is consistent with previously acquired beliefs and attitudes. Bergstrom, Moehlmann and Boyer (2006) believe that children easily acquire information that dove-tails with prior knowledge (e.g. the results of parental instruction) but when it clashes then source evaluation becomes important. For example, when told the colour of a ball they had previously seen, children did not accept false information from either reliable or unreliable sources. But when they had not seen the ball they accepted the reliable source over the unreliable one. Source evaluation will be important if the child does not possess first-hand knowledge of the information in question. Koenig et al. (2004) show that children are able to remember sources of information as well as which sources to trust. When confronted with two speakers giving nominal labels to familiar objects children learned which one to trust and then accepted that speaker's information when novel objects were named.

So, children show selective trust when they don't possess first hand knowledge of the information in question. How is this happening? How do children navigate complex information environments with ease?

In order to obtain useful information the agent must pay evaluation costs. These are the costs of discerning good and bad information, or what is likely to be good and likely to be bad information. The agent has two sources of information cues available to him, *content* and *non-content* cues. With the children just discussed, accepting information because it dovetails with existing knowledge is content mediated, accepting information because the source is trusted is using a context cue.

Richerson and Boyd suggest that the empirical evidence leads us to believe that there are psychological mechanisms whose purpose it is to recognize these information cues in a complex world and infer the best course of action. These authors postulate a conformity bias and a prestige bias on the basis of ethnographic evidence (e.g. Henrich and Gil-White 2001) and computer models (e.g. Henrich and Boyd 1998). They favour a dominant role for non-content biases.

2.6

Minds respond to information complexity with biases

Biases in this sense are psychological mechanisms favouring the uptake of certain information from the environment over other kinds of information. Biases, as I will explain, may latch on to intrinsic features of the information itself, or features extrinsic to the information, such as its source. I suggest that biases are essential in informationally complex environments and their emergence is a necessary result of the beginnings of cultural information transfer because of two main factors:

1. Biases help to tractably solve the problem of how to get the best information possible given a multitude of possibilities.
2. Biases can help agents to avoid the worst information. This is especially important once defectors exist (who take advantage of such things as gullibility, or who copy-only without testing information by trial and error). At some level biases may operate in analogous ways to antivirus software or firewalls and could shield the mind from information that has a higher chance of being dangerous. In a world of varying memes, child development (as the result of genetics or social learning) must install these measures to protect the growing agent. This is especially so in the face of 'killer' memes that have no interest in agent fitness.

I will argue in chapters four through six that memes exist in the strict sense. If memes are important in human cultural evolution then it should be apparent that biases will come to shape the flow of memes between agents as agent fitness wrestles with the memetic fitness of copied cultural traits. As soon as there were memes, minds were forced to evolve biases.

Henrich, Boyd and Richerson favour agent dependence on non-content (or context) biases in acquiring information as a cheap reliable solution to the problems of individual adaptive choice. I will now discuss their two most frequently cited non-content biases. These are conformity and prestige-biased copying. I will conclude that these biases may well exist, but their adaptive value will vary greatly across a range of circumstances, therefore, they are not the full story of how we consistently navigate complex information environments. More will be needed to explain human culture, so I will then go on to discuss content cues and the effects of group density and group structure on cultural evolution.

Two Kinds of Non-Content Cues

It is important to understand the nature of these evolved psychological mechanisms and how they function in shaping the flow of information between humans. Imagine a hypothetical agent struggling to obtain fitness information. This agent has a choice of behaviours (p,q,r,s) in an opaque world. The fitness value of the behaviours on offer are arbitrarily $p > q > r > s$. Importantly, s leads to death. It seems obvious that the priority for the agent is to avoid death and only then to select information that will lead to fitter than average behaviour. The agent must first avoid the worst information. But fitness information is not available to individuals, so we must track *fitness correlates*.

1. Conformity and Gullibility

A conformity bias leads the agent to tend to copy cultural variants that are more common than other variants. Gullibility is a version of a conformity bias, where an agent adopts the behaviour other individuals tell him to. I will now explain how a conformity bias is better than copying at random, and that it can reduce transmission error. But, conforming eliminates optimal information, corrupts potentially fit vertically transmitted information, and may actually be the result of simple familiarity processes rather than a genetic disposition to conform.

Overall agents are trying to decide how to act in the future, but predictions about the future are problematic. The number of variables to account for is too large. But agents can use reliable cues to help their decisions. For example, an important fact about populations (assuming inheritance) is that the statistical distribution of traits at time (t) dictates the statistical distribution of traits at time (t + 1) to a greater or lesser degree. This information can be used by the agent to infer what traits were common (and therefore probably adaptive) in the past, and can aid the decision as to which traits to copy. Copying a common trait would seem to be more adaptive than copying at random. Conforming is therefore tracking a fitness correlate, in this case the frequency of a trait. If it's common it's good.

It seems conformity focuses imitation upon adaptive behaviours. This is not the only theoretical advantage of conformity however. Conforming will also tend to reduce transmission error and avoid dangerous information. Transmission error is reduced as follows. Suppose information is noisy. Each individual has a 60 per cent chance of noticing that behaviour A is adaptive (perhaps by individually comparing the hunt return of blow guns versus bows), so by using individual learning an agent gets it right with a probability of 0.6. However, if the agent samples 10 models and adopts the majority behaviour then he has a 0.75 probability of getting the adaptive behaviour A, assuming

that everyone else had a 0.6 chance of learning it on their own as well (Henrich and McElreath 2007). However, a conformity bias will not obtain *maximally* fit information (perhaps spears are best, but only one tribesman uses them) and it also runs the risk of eliminating variation and with it useful novelties. Therefore, a purely conformist system will become stagnant. It is interesting to note that there are many human systems in which copy-only or conformist strategies are punished. One example would be plagiarism in academic writing. Clearly a plagiarize-only strategy would lead to stagnation, so we need to encourage individual variation to some degree. What is good for the individual (plagiarizing) is not necessarily what is good for the group (diversity). So, such high-risk, high-reward systems are going to be adaptive at the level of the group.

Mathematical models predict that learners ought to prefer conformist transmission to vertical transmission. However, I again stress the need for caution in interpreting the mathematical models. There are some reasons why vertical transmission may hold more influence than simple models suggest. Firstly, vertical transmission is likely to be important because parents are often the *only* models that infants have for many behaviours. Infants spend far more time with parents than with any other potential model. Secondly, there is environmental variation *within* the group too. What is good for the parents is likely to be good for the kids because they inhabit the same environment. Thirdly, cultural traits may influence which subsequent cultural traits are learned. Often culture exists as suites of behaviours and beliefs. If an infant learns part of a suite from his parents then he may more easily pick up other facets of that suite (think of religion). Fourthly, children will be a nearer genetic match to their parents than to other members of the group. So, for any traits that depend upon particular phenotypes in order to successfully implement them, there is more chance of a neat gene-culture 'fit' between offspring and their parents. Finally, there may be coordination issues. If there are norms expected or enforced in particular families then children will be

constrained in their learning. All these factors need to be incorporated into cultural modelling. It is true that learning continues long after we cease to be infants, but what is learned as an infant/child can massively bias what can subsequently be learned by the agent. All of this is going to make vertical transmission more important than conformity in some cases.

A last consideration with any proposed psychological bias is its biological plausibility. I wish to suggest one mechanism via which the brain itself may be predisposed to conformist learning. Then, donning a skeptical cap, I will suggest the predisposition may actually be toward familiarity and this could suffice for what appears to be conformist behaviour and also explain variation in the tendencies of different people to conform.

Humans (and great apes) possess a set of pre-motor cortical neurons which activate in response to the observed actions of other agents. These neurons are the same that fire when the observer enacts the task observed (Ramachandran 2000, Miall 2003, Byrne 2005). Such mirror neuron systems would no doubt be active when observing cultural practices. We would expect sets of neurons to fire in frequency-dependent ways according to which cultural practices are most common in a population. Assuming some sort of Hebbian learning[3] in these neuron populations, where synapses that are active more often strengthen, then those neurons that fire most frequently will 'wire' most strongly. It may then be the case that attempts to imitate frequently observed behaviours have more chance of success than attempts to imitate seldom seen behaviours thanks to this priming. Basically it may simply be neurophysiologically *easier* to conform. This mechanism could be selected for because common practices (in many situations) will not be maladaptive and indeed may be fitter than other solutions.

---

[3] Hebbian learning is where an axon of one cell excites another cell, or repeatedly takes part in firing it, and then the first cell's efficacy as a trigger for the second cell's firing is increased. If a presynaptic action potential occurs a few milliseconds before or during the postsynaptic one then the efficacy of the synapse is increased in lasting fashion (Keysers and Perrett 2004).

Chaminade (forthcoming) further argues that mirror-like properties are found with neurons in other regions of cortex too, such as sensory neurons and emotion controlling areas. The term 'resonance' has been introduced to refer to these more distributed phenomena. It appears that producing an action or sensation similar to an observed action or sensation is 'a prepotent response that is actually inhibited in order to execute the correct response to the cue', (Chaminade, forthcoming, pg 15). Furthermore resonance seems to occur more strongly when agents observe an action with the intention of imitating it. Finally, resonance may arise at least in part by associative learning when an agent performs an action and observes himself performing the action (this could explain the bias to finding mirror neurons associated with grasping actions).

So it seems that there may be a basis for *frequently observed* actions becoming *easier* to execute due to resonance and Hebbian processes.

But now we must distinguish *familiarity* from *conformity*. It is possible that an agent is most *familiar* with (observes most frequently) a practice that is not what the majority are doing. This could be because his mother or teacher practices this behaviour and he sees it over and over again. If the agent were to copy this behaviour he would be practicing a familiarity bias (a *content* bias) rather than a conformity bias. Unconscious familiarity bias, therefore, even in the very young remains biologically plausible thanks to association, Hebbian mechanisms, and resonance. A familiarity bias may *look like* a conformity bias. But there would be important differences, conformity is dependent on context. Familiarity bias according to this mechanism would merely result from selection for the fundamental brain processes of Hebbian wiring. It may *appear* that we practice conformity, but we need to clearly separate this from familiarity processes when studying these phenomena.

The degree of conformity varies on an individual basis (Asch 1955, Henrich and McElreath 2007, pg 564). This needs explaining, and the preceding discussion may go someway to accounting for this fact. It may be that there are different sorts of conformity bias and they are not all the results of one process.

Overall we can be skeptical of the standard account of conformity bias. We can challenge the value of conformity because of contexts in which it is maladaptive, and we could challenge the conformity bias' ontology on the basis of an associatively facilitated familiarity bias. What appears to be conformity may actually be based on a familiarity bias. This would explain apparent 'degrees' of conformity among different people. It would also mean that actual transmission patterns will vary somewhat from those predicted by pure conformity theory. Agents will conform but not when they are more familiar with alternative procedures.

What this analysis of conformity has shown is that although conformity can focus learning on fit behaviours and reduce transmission error, it is unlikely that conformity (even coupled with the prestige processes that I discuss next) is sufficient to explain human culture consistently across many contexts. There are likely to be other biases that help to preserve information that would otherwise be lost through conformist processes. There are also likely to be social and population structures in place that help to channel cultural transmission. And finally, the mathematical models that continue to suggest that conformity is adaptive in a wide range of social scenarios need to grant more weight to the likely importance of vertical transmission.

2. Prestige Processes

Henrich and Gil-White (2001) argue for the existence of prestige processes as an emergent product of psychological adaptations that evolved to improve the quality of information acquired via cultural transmission. We must distinguish

prestige from dominance. In prestige processes there is freely conferred deference to the model. Copiers will grovel or reward the model in order to gain access to copying him. In dominance hierarchies the dominant individual is feared and imposes his control by threat of (or actual) force. We must also distinguish success and skill-based biases. In success-biased copying the agent copies the individual who shows the objective, often quantitative, mark of success in a particular domain, for example the hunter who brings back the most kills. This can be difficult to assess for tasks whose results vary from instance to instance. The agent could also copy according to skill, this is a qualitative method. A carpenter who often cuts himself and breaks his tools may seem like an unskillful model without even assessing his products. In prestige-biased copying, on the other hand, the agent looks for cues of deference and copies the individual who is the target of these deference behaviours. Deference markers are indirect measures of special expertise and may include: being gazed at more, being excused from social duties, receiving gifts, being helped, or generally being 'kissed up' to. These deference behaviours are freely conferred.

In a complex information environment sycophantism marks successful models. Having a preference for popular models tends to get good information for the individual agent. This is because other agents have already evaluated the skill or success of the prestigious individual and started copying him.

Furthermore, according to Henrich, Boyd and Richerson, prestige bias can inject the required environmental tracking into conformist populations. Basically this is a way for social groups to overcome the limitations of a purely conformist strategy that tends to stagnation. Groups can preserve adaptive novelty if a prestigious person invents it. However, as with conformity, there are two important issues here. (1) We must ask, do prestige-biased processes actually exist, and (2) are they really that adaptive? Having called the

mechanism of the conformity bias into question, I will now do the same for prestige bias.

In answer to (1): It is certainly possible in small social groups for everybody to learn from a single prestigious source. Indeed, evidence suggests that even in larger groups people tend to copy *locally* prestigious individuals. For example, Rogers (1995) writes on the diffusion of innovations. He notes that people adopt practices after a subjective assessment of practices that others similar to them are using. The individuals generally modelled are locally high in social status, well connected, and respected. But are such processes really 'prestige'-based? Are we seeing freely conferred deference in exchange for information goods at an individual level, or is this perhaps a version of a conformity bias at work? Individuals are just copying who everyone else is copying locally. There may be some sort of second-order conformity rule in place. Furthermore, are some 'prestige processes' scaffolded by institutional facts rather than psychological mechanisms? Agents are born into a social world that is structured in such a way that certain individuals are role occupants of highly visible roles. Children are taught that it is good to copy these people. These people may occupy the roles because they have demonstrated objectively measured proficiency in a particular task (think of Roger Federer). We copy those who social institutions have tagged (or made highly visible) for us to copy, without the need for an internal prestige bias. Or, we end up copying successful individuals merely on the strength of a conformity bias. One must not be too quick to attribute phenomena to genetic adaptations, or to cognitive biases that favour prestige copying. Proponents of processes like prestige biases must fuse these purported mechanisms with the construction of social environments that direct learning and highlight the targets of learning. Throwing doubt over the exact mechanism, however, does not deny that there is a pattern of information being copied from certain socially marked individuals. What is in doubt is that internal prestige biases are important in this process. Other biases or social institutions could suffice.

(2) I now note that employing a prestige bias isn't as adaptive as it first seems. There are serious practical problems for agents following a prestige bias. Prestige bias works when prestige is a true proxy (or fitness correlate) for success. If an agent wants to copy the successful then they want to be sure that the target actually is successful. In the past material wealth could be seen as an honest (hard to fake) signal of success. This is not a cue of prestige, of course, but material wealth can buy deference and then deference can act as a signal of prestige thereby obtaining more deference. However, easily available credit, for example, could undermine the honesty of the material wealth signal. Furthermore, it may be that a prestigious individual is successful *despite* some behavioural traits, or possesses useful traits that require other things such as the right genes to be useful behaviours. Learners need to compare themselves to the model and decide if they are similar enough that copying will work. This may lead to preference for copying those who share ethnic markers (Henrich and McElreath 2007). Indeed, ethnicity demarcation and ethnocentric behaviour seems to be predicted by the necessity for biased social transmission in a complex information environment. Copying too freely all the behaviours of a prestigious individual can lead an agent astray. Copying needs to be restricted to those targets who share a niche with the learner.

So prestige bias may not be as necessary or prevalent as Henrich and Gil-White argue and prestige processes are not always clearly adaptive.

Prestige, however, can inject the required variation into an evolving cultural system that is based on variation-eliminating conformist tendencies. Even though there are costs in copying prestigious individuals, these may be paid for by group selection for prestige hierarchies and institutions, which on average raise the fitness of the group as a whole.

I have now discussed two non-content biases, conformity and prestige, which may aid agents in navigating the information environment *if certain circumstances exist*. Though they seem to have theoretical foundations and they seem to be extant, the exact nature of these biases is controversial. Controversial, too, is the proportion of their contribution to sustaining complex adaptive culture. I doubt that these two biases alone can explain the accumulation of complex adaptive culture. I also doubt that these biases are genetic adaptations and I will argue for this claim further in the next chapter. More is needed to explain the state of human culture than conformity and prestige biases. I have introduced several other possible biases such as ethnic identification, gullibility, skill, and success biases, all of which may have a significant role to play. I have also introduced the possibility of familiarity, a *content* bias. I turn now to arguing for the importance of content cues in acquiring adaptive culture. After that I will argue that certain types of social organization, such as large effective population sizes, aggregating mechanisms, and certain institutional facts are also important in retaining adaptive culture.

Basically, populations of agents that are too rigidly locked into conformity and prestige biases will do poorly. These context biases need to be developmentally plastic in order to be able to respond to a variety of social situations.

Content Cues are Important

Henrich, Richerson and Boyd's models favour domain-general 'source-biases' for information acquisition. In their view prestige plus conformity together works to generate adaptive novelty and sustain it. Content takes a back seat. Sperber, on the other hand, favours content, 'all learning is biased by content' (Sperber and Claidiere 2006). Sperber alludes in several places to familiarity (an agent-internal measure) and to the *transformation* of inputs to fit familiar cognitive representations. Sperber also discusses ecological forces (for

example the ingredients available for a soup in a particular region) as being important non-bias forces in cultural transmission. Human niche construction and our choice of environments will affect what information content can persist. A soup recipe that uses pumpkin is not going to spread through a population that grows potatoes.

Content biases may be either reliably developing products of our species-shared genetic heritage, for example the tendency to prefer salt to chalk on pop-corn, thanks to our taste receptor systems (Henrich and McElreath 2007), or they may be culture specific, or even individual specific (as what one has already learned biases what is likely to be learned next). Content analysis is a difficult task. How could a naïve agent possibly discover that boiling water prevents many water-borne illnesses when the causal agents are not visible and the results of drinking unboiled water are unpredictable? In addition to the non-content biases discussed above, the following sorts of factors all determine whether a piece of cultural information is taken up by a given mind.

1. The complexity of the information. A one-step behaviour is going to be easier to remember than a three part one. A story is easier to remember than a list.
2. The ease of acquisition (familiar information being easier to assimilate than new).
3. The appeal of the information. Does it look important? Does it correlate with important information already held? Does it appeal to what the agent values?
4. The visibility or attention grabbing nature of the information.
5. The 'fit' between the new idea and ones already held.

Cultural traits that are attractive to psychological content biases spread widely.

Eckel and Wilson (2007) demonstrate the power of content over non-content cues in some situations. They examine a 3x3 coordination game. In this game players must choose one of three strategies: (1) a safe low-risk, low pay-off option, (2) a riskier but overall more efficient option, or (3) a third option which is dominated in all outcomes. When the game is played there are two possible outcomes where players *and* their opponents can both win. This occurs if both choose option one or both choose option two. For example if both players choose the safe strategy (1), then they both win. If both choose the riskier strategy (2), then they both win. Asymmetric strategies may result in one or the other winning but not both. However, strategy (3) always loses even when both players play it.

After each round all players are informed as to the strategies of either a simulated or real player, who has either high status (royal) or low (commoner). Status is determined by administering a quiz at the start of the experiment and quiz scores are divulged. The game is then played again. These experimenters found that players are easily encouraged to switch to strategy (1) by observing the revealed strategy and it did not matter whether the player of the revealed strategy was a royal or a commoner. However, if the revealed strategy was the riskier option (2) then the play of this commonly observed player had a positive impact on the subsequent play of strategy (2) only in case the commonly observed player was of high status. But even a royal player could not move other players to play strategy (3).

The conclusions drawn from this experiment (and others like it conducted by the same researchers) are that the prestigious individual can influence choice between two equilibria, one of which has high-risk high-payoff and one which is less risky but pays less. But the high-status individual cannot move subjects toward a strategy that is dominated on all dimensions (3). What these sorts of experiments seem to show is that prestige-type processes can shunt agents to a more efficient strategy but if the information itself has little appeal (i.e. is

dominated by all strategies) then prestige can't move an agent toward it. What we can conclude is that theories like those of Richerson and Boyd that emphasise non-content biases need to be fused with content-based approaches to the transmission of cultural information. Content is important, especially when it is very unattractive or very appealing.

I will now move on to discuss the structure of social groups and information pooling mechanisms. Alongside the biases (non-content and content), these features of cultural groups significantly shape cultural processes.

2.7 Crowd cognition and social group structure

There are features of cultural group structure that influence the acquisition and retention of adaptive culture. I will now explain that a complete theory of cultural evolution needs to take into account group size, information pooling mechanisms that the group possesses, the degree of connectedness of the group to other groups, and the variety of cognitive devices within a population. Differences in these features in specific cultural groups will have effects on the nature of cultural evolution in these different groups. Different patterns of cultural evolution will have the effect of selectively favouring different psychological biases over others in different circumstances. Because cultural circumstances can change so rapidly, this suggests that cultural transmission biases need to be developmentally flexible.

The size of a cultural group will affect the success the group has in navigating a complex information environment. Biases may not take effect until threshold group sizes are met. For example conformity increases with group size (Bond and Smith 1996). But, notably, crowds are good at collectively solving cognition problems (Surowiecki 2004, List 2005). There is an important difference between conforming to the mode (conformity bias) and averaging everybody (crowd cognition). *If* agents are able to average a large sample of

potential models then they will obtain, on average, the very best information there is.

Surowiecki describes Galton's observations of how a crowd of naïve villagers can pick the exact weight of an ox by each guessing its weight and then writing it down on a piece of paper (as entries in a competition). Once these diverse, independent, decentralized decisions are made then one can aggregate the information by averaging the crowds' guesses. Such results are surprisingly accurate across a range of tasks. Another example followed the explosion of the Space Shuttle Challenger. Several companies all manufactured parts for the shuttle and initially it was unknown what had malfunctioned. Immediately following the disaster the shares of all these companies fell, but within hours all had stabilized except for the shares of Morton Thiokol which continued to fall, twelve per cent in one day. It turned out that Morton Thiokol made the O-rings that were subsequently proven faulty. The traders had no way of knowing this for sure at the time of the explosion, but pieces of information distributed in the market had aggregated to the correct decision. It is still possible that some one piece of information may have spread like wildfire through the investors and led to the stabilization of the other stocks, perhaps we do need more information for any one instance of alleged crowd cognition. But there are many other examples of the phenomena in futures markets, in sports predictions, and so on that it is hard to deny that the phenomenon occurs.

However, there are important caveats to the success of this phenomenon. The *method* by which individually held information is pooled into a useful body of knowledge is crucial. This *aggregation procedure* must be a good one because it is possible for the process to be lossy with respect to important information. Here is an example of such lossiness. There exist cases known as 'doctrinal paradoxes' where a majority view on each of a set of premises does not carry through to a majority view on the conclusion if only the conclusions are

aggregated[4]. There needs to be independent aggregation of each premise. However, provided the crowd is diverse, the agents independent, there is decentralized decision making, and there is a method of aggregating the group decision, then crowds' performance in certain types of task surpasses (on average) that of experts.

Crowd cognition works because if lots of people possess at least some piece of accurate information about the world then aggregating all those pieces of information (and subtracting error, which is easy if everyone makes different errors) leads to the best decisions.

Henrich and Boyd (2002, pg 109) make a parallel point when they note that their models show that adaptive cultural evolution can occur, 'even when representations are continuous and inferences are biased against adaptation'. This is because poor individual inferences can be compensated for by larger pools of social learners and the fact that people don't all make the same error.

Crowd solutions to cognitive tasks are exceptional when the options are discrete or quantitative. Of course this method is corrupted when there is extensive conformity or prestige-biased copying occurring as the average of the crowd then approaches the behaviour of prestigious individuals or that of the mode. The strategy relies on the different judgments pre-aggregation being as independent as possible.

---

[4] For example, suppose three judges must judge on whether a defendant is liable for a breach of contract. According to the law, the defendant is liable (proposition R) if and only if he performed X (proposition P) whilst having a contractual obligation not to do X (proposition Q). Thus legal doctrine requires R *iff* ($P\&Q$). If all three judges accept R iff ($P\&Q$), and judge 1 accepts both P and Q and, by implication, R. And judges 2 and 3 each accept only one of P or Q and, by implication, they both reject R. Then if the court applies majority voting on each proposition (including on R iff ($P\&Q$)), it faces a paradoxical outcome. A majority accepts P, a majority accepts Q, a majority (unanimity) accepts R iff ($P\&Q$), and yet a majority rejects R.

So what we can conclude is that to get the best solutions there needs to be group suppression of some conformity and prestige processes and cognitive diversity must be actively selected for. This is because we need diversity in the options *and* in the decision makers. This is the formula for obtaining the best adaptive solutions. These solutions can then spread by mechanisms like prestige-biased transmission once it becomes obvious that they are successful.

Furthermore, people who tend to hold minority views can lead majorities to more systematic processing of social information. If the minority view is coherent and consistent then the majority view gets examined closely for failings. It is good for a group to have some minority dissenters. Again what is optimal for group fitness may not be optimal for any one individual. If there is group selection then cultural transmission models need to be reassessed.

Therefore, if we believe that cognitive diversity is selected for, then we must forgo theories that uphold universal biases. Conformity and prestige, even together are non-optimal. Groups use other mechanisms of information appraisal and models of cultural evolution must reflect this as it will affect the transmission dynamics. It would be much more adaptive for a *group* to have a bunch of different, plastic minds than an array of identical biases.

Prestige processes could possibly be seen as an example of crowd cognition. By looking at the *distribution* of deference in a population, an agent can aggregate the information that everyone else has extracted from the system in choosing who to copy from. Such crowd solutions will often be better, on average, than any one individual's decision is, even that of an expert.

The overall point here is that populations that interact in the right way and populations that consist of diverse individuals are likely to happen upon good adaptive solutions. Indeed, non-content biases may hinder these crowd cognition processes.

Conformity and prestige processes are often good for agents. But we must also allow for the effects of crowd cognition. Diversity of cognitive devices would appear to be optimal for a cultural group, so the biases are unlikely to be universal because individual fitness is wrestling with group fitness.

Group Interconnectedness is Important

Thinking about between-group dynamics may also help to explain the variation we see in the *rate* of cultural evolution and adaptation in social groups (see Diamond 1997). The rate of adaptive evolution depends on the natural logarithm of the *effective* population size (Henrich and McElreath 2007). Basically, if the effective population size is large then the rate of adaptive evolution is fast. *Effective* population size takes into account migration and information trade between spatially or culturally separated groups, not just the size of one particular social group. Populations that consist of a network of groups or 'nodes' that are isolated and do not trade information evolve cultural adaptations more slowly than populations where nodal clusters are connected in a densely-knit or glocalized network (see Boase and Wellman 2001). Ramified networks, where nodes are sparsely connected to immediate neighbours would evolve adaptations at an intermediate rate (figure 2.4). Basically, the more you talk to your neighbours and the more accessible nearby groups are, the more easily you will acquire adaptive information.

The paradigm case of a node becoming disconnected from a network is Tasmania between 12,000 and 10,000 years ago when it was cut off from the Australian mainland by rising seas. What followed was a dramatic regress in cultural complexity and tool mastery in the Tasmanian population (see Henrich and McElreath 2007 for discussion). This was a direct result of the lack of information trade in and out of the Tasmanian node of the network. This

reduced the effective size of the Tasmanian group below that required to sustain some complex cultural adaptations.

So models need to account not only for conformity and prestige-type processes, and not only do we need to model the importance of vertical transmission, but we need to add crowd aggregation, diversity selection, and group connectedness variables to models of cultural transmission and evolution. Because of this array of non-static variables, biases are unlikely to be universal because different variable settings in different contexts will lead to selection for different arrays of cognitive biases in different populations. I will argue in the next chapter that this is consistent with neurobiological and learning theoretic constraints to cognitive development.

Figure 2.4

Three Models of Network Structure

In a densely knit network severing one information channel has little effect on the flow of important information between the nodes because there are still many inter-nodal connections to all nodes. In a ramified network cutting one channel can isolate nodes and the availability of information can decay. If nodes represent populations then we see that cutting one link in a ramified network of populations can isolate a group from all other populations. Inter-group information flow ceases and the quality of the isolated group's information can decay. This is what happened when Tasmania was cut off from mainland Australia by rising seas. Cultural

2.8

Recap of humans' response to information complexity

I have shown that social learning emerged at some point in our lineage and that our truly imitative capacities led to an explosion of culture. But the existence of culture has necessitated that we evolve mechanisms for harvesting adaptive information from the cultural world rather than maladaptive information. Psychological biases have emerged in various guises that depend on context or content, we have created institutions that teach us what to copy, and there are structural features of groups that sustain adaptive culture. However, there is an ever present danger of defectors exploiting our biases. Think of contemporary mass-marketing and conformity. In such cases a conformity bias may not protect the individual agent from exploitation because the message is conveyed to the entire population simultaneously. So there are social situations in which it is adaptive to conform and others where it is not. These are defector (or indeed predator) problems in that the groups' resources are flowing to a select few individuals. This is the case with other biases too. There are situations in which they are adaptive and situations where they are not. It would, therefore, seem that biases, or our capacity to switch between or form new biases, need to be *flexible* in order to adapt quickly to changes in the information distribution in a culture. At least some part of our biases must be responsive to the environment, perhaps even encoded by particular social structures that are responsive to change.

Furthermore, it is likely that the important vectors in cultural transmission have not been static over evolutionary time. In static environments, vertical transmission may have been sufficient for cumulative cultural adaptation (Mameli 2006). But horizontal transmission is important in other circumstances such as in more contemporary environments where the knowledge of the previous generation is outdated.

In the Pleistocene with smaller, genetically related groups, the biases may have been quite different to those we see today. Richerson and Boyd think that vertical transmission is less significant than other factors in cultural evolution. This may be the case today where individuals are assailed by information horizontally and internal world-models differ from generation to generation, but seems unlikely to be the case in our environment of evolutionary adaptation where children learned skills from same-sex parents and elders. It seems that there have been shifts in the forces responsible for the net flow of social information across the homo lineage's time span. I suggest that the focus upon the forces involved needs to be clarified. We need to address the forces involved at a *particular time*, in a *particular context*. These are macro-evolutionary considerations (Richerson and Boyd 2005). As Mameli rightly points out, knowing how biases operate today doesn't necessarily tell us about the way they operated in pre-industrial, let alone pre-agricultural societies. I suggest that the composition and structure of such societies, being dramatically different from the societies in which a lot of bias data has been collected, suggests that we ought to be very careful in generalizing from data to rules. We may even ask if the cognitive biases even existed in the past. I suggest that the following transitions are likely to have occurred.

1. The primitive form of cultural transmission was probably simple within-generation flow. This is because the kind of behaviours transmitted were of immediate significance such as food calls. Subsequently, the Homo lineage seems to have been selected for its vertical and oblique transmission capabilities (Hewlett and Cavalli-Sforza 1986). Culture enhanced the ability to track the environment. Culture then provided a selective pressure for communication, for increased group size, and increased inter-group complexity. The cultural environment has changed rapidly over historical time driving a tendency toward more complex and extensive horizontal transmission in recent ages as vertical information has become less closely matched

to a changing environment, and as the diversity of possible behaviours increases. This may have approached extremes in contemporary society where the environment of today's youth, an environment of credit, service industries, virtual reality, global warming and freely available information online, is massively different from the environment of their parents. In such an environment horizontal transmission and importantly individual learning ought to be significant (these are cases where u approaches one in the formal model). It may be suggested that we ought to study cultural or proto-cultural transmission in primates in order to theorize about the nature of information transmission in our ancestors. But if the marker of human uniqueness is the ability to imitate and transmit adaptive behaviour vertically and culturally, then this may merely confound the issue.

2. The transitions detailed in (1) entail a switch from predominantly high fidelity copying of kin (once vertical transmission was important) to including at least some averaging of peers as this ought to obtain the best information possible in contemporary environments.

3. Epistemic technologies such as teaching institutions and public resources increase in availability over time. This permits complex cultural information to be reliably transmitted to the relevant individuals. Institutions flag traits that are to be copied and this helps to save on assessment and acquisition costs as cultural worlds become complex.

4. The more culture there is, the more important that cultural group selection becomes because culture tends to make cultural groups more internally homogenous and more different from other groups over time (Durham 1991). If biases cause us to *conform* and *cooperate* with those like us and diverge from other cultural groups then human groups may approach the 'super-organism' status that Sloan-Wilson details (2007). Selection moves up a level in the hierarchical Darwinian model.

At this point I must highlight the importance of cultural group selection. Mameli (2006) argues that the biases identified by Richerson and Boyd have costs. For example the conformity bias eliminates useful novelty. Success biases impose groveling costs on copiers and letting out secrets costs on models. There is also a risk of copying behaviours that are not adaptive merely because they are displayed by prestigious individuals. Sterelny (2006b) argues that these costs may have been paid by group selection. Group selection is not an issue discussed at length in this thesis, but I recognize that a defense of cultural group selection is central to several themes I raise in this and other sections.

The point that (1) – (4) illustrate is that the vector components of cultural flow are not static over time. This is important because as cultural forces change then minds must be able to quickly adapt. This would seem a *prima facie* argument against the full genetic encoding of psychological biases. I turn to this issue of what the underlying nature of the biases is in the next chapter.

2.9

Summary

So we see that once imitation arises then culture accumulates and information environments become increasingly complex. Agents need to employ rules in order to acquire useful information. As well as learned individual heuristics, such as, 'do what worked for me in the past', these rules include non-content and content biases. My central point is that there is a dynamic flow of information between humans and that *some kind* of heuristics and biases influence this flow in adaptive ways. But what do we make of the nature of these biases? For importantly, as Driscoll notes (2008, pg 105), 'the use of acquisition decision rules only solves the problem of adaptive individual choice if acquiring acquisition rules and using them to acquire fit behaviour is itself a tractable learning problem.'

I will argue in the next chapter that such rules do not need to be genetically specified in order for this constraint to be met. Indeed, if, as Driscoll argues, we have a significant socially transmitted element to our biases, then biases such as conformity are plastic and could even be recent evolutionary adaptations. Conformity as a rule could conceivably have developed in response to recent increases in group density. In small groups it would be possible for everyone to access and copy a successful individual. Matt Ridley (2003) notes that humans invaded every viable niche on earth and then were forced to increase in density. This crowding may be responsible for an increase in the quality of information and may have bred a host of biases.

The main conclusions from this chapter are twofold. First, there are issues of methodology in the study of culture. We can use the conceptual logic of evolution by natural selection and we can mathematically model states of affairs. Mathematical models reveal novel insights but currently lack the required resolution to be sure of their results. Culture is a complex system and such non-linear systems exhibit interesting dynamics. If the origin states of two models are arbitrarily close the result can be two vastly different conclusions. Therefore, it is important to ensure that all *important* variables are accounted for in such models. This is not yet the case with respect to models of cultural evolution.

The second main conclusion here is one of paradigm consideration. There are current assumptions in the literature that the psychological biases are universal and genetically encoded. The current analysis suggests that neither of these is likely to be the case. The success of the biases is context-dependent, so genetic encoding could lead to terrible maladaptation without the ability to rapidly amend the situation.

Chapter 3: Neural Constructivism and the Psychological Biases

Overview

I have just shown that cognitive biases of some sort must exist in order for human agents to usefully navigate complex information environments. I have described some of the biases (conformity, prestige) that appear to exist according to ethnographic data. There are likely to be many more. It is now time to consider the nature of these biases. I will argue here that high-level, domain-general cognitive biases are not likely to be the products of genetic evolution. I will begin by noting that cultural beliefs vary and so, too, might cultural cognition (3.1). Then (3.2) I argue that it is far from obvious that cognitive transmission biases are genetic adaptations. In order to examine how these biases might arise, I introduce a tracking constraint on the nature of our cultural cognition (3.3), and then neurobiological and learning theoretic constraints on the development of these biases (3.4). I also argue that the nativist replies to those who see human cortex as an organ of unspecified plasticity fail (3.5). I will describe the paradox that results if we reject nativism while accepting the learning constraint and offer a solution to this paradox that satisfies the tracking, neurobiological, and learning theoretic constraints (3.6). I will conclude by detailing the predictions that this view of mind makes and offer empirical support for these predictions (3.7). I will then conclude that the evolutionary psychology picture of a massively modular mind, where biases are genetically coded modules, is wrong. I suggest that minds are evolved for their plasticity (3.8) and as a result cross-cultural differences in the cultural and physical niche determine to a significant degree what cognitive apparatus an agent is endowed with. I suggest that this is likely to apply to the transmission biases discussed in the previous chapter too. This is consistent with my suggestions that transmission biases are not always adaptive in all contexts. The function of the mind is, to a degree, dependent on the cultural milieu, and the cultural milieu in turn depends on the function of the mind. The nature of

cultural evolution effectively evolves culturally. There is, therefore, no reason to believe that the cognitive biases shaping the flow of cultural information are fixed. Culture undergoes transitions in its dynamic flow over relatively short (genetic) time spans and culturally relevant cognition must therefore be fluid.

3.1

Cultural beliefs vary, so might cognition

It is certainly clear that beliefs can vary from culture to culture, situation to situation. One may believe that Jesus is the source of good information and so choose to accept as true or useful anything attributed to Christ. Alternatively one may do whatever the majority of a crowd around one is doing. This may also be belief-based and conscious. For example it's possible that we are thinking, 'mum always told me to do what everyone else is doing', as we conform. But it is also possible that we conform in this situation because of an unconscious cognitive mechanism. Such a mechanism could be psychologically primitive or 'innate', or it could be the result of experience and propositional knowledge. Such a mechanism may guide our behaviour by taking situational inputs and using them to direct attention for example. This sort of mechanism might be called a piece of cognitive architecture, or a learning mechanism for guiding behaviour.

It is physically possible that learning mechanisms, like the number of limbs we have, are in place at birth, but it is also physically possible that the mechanisms are causally effected by development and emerge uniquely for each developing agent, like the pattern of blood vessels on the retina. A final possibility is that propositional knowledge, through association with action plans, becomes hard-wired by habit. We may begin consciously doing what Jesus tells us to do, but eventually we don't even notice we are doing it and the bias is driven by non-conscious processes, the habit has become unreflective. This would then be an

architectural change rather than merely addition of information to a database to be called on in rational deliberation.

Propositional knowledge (belief and so forth) is contingent and learnable. Differences across cultures here are called differences in cross-cultural *beliefs*. But there could also be differences in the mechanisms or architecture of the learning devices, such differences would be differences in cross-cultural *cognition*. I will argue in what follows that the mind/brain is plastic in the right sorts of ways that the mind/brain's own 'learning' of the cognitive architecture which will instruct it in acquiring propositions and behaviours is also contingent on circumstance. To use a tired computer metaphor of the mind, these differences would be differences in hardware or firmware, rather than differences in software or databases.

3.2
Biases are assumed to be genetic adaptations

I will now outline the standard position regarding cultural transmission biases. That is that they are evolved adaptations that are likely to be genetically coded. I will point out that this is an unsupported assumption.

Henrich and Boyd (1998), Boyd and Richerson (1985), and Richerson and Boyd (2005) argue that the flow of information between human beings is biased by internal psychological mechanisms. They provide evidence (e.g. Gigerenzer and Todd 1999, Gigerenzer and Selton 2001) for the existence of biases including conformity and prestige biases, and speculate that as yet unidentified biases are likely to have evolved also. Their reasoning runs as follows:

1. Optimality models suggest that bias X is the optimal strategy in environment Y (computer simulations)

2. Bias X exists in a variety of cultural groups (psychological and behavioural data)

3. Therefore, natural selection has selected for bias X

A bias in this sense is a decision making strategy employed by an individual agent. A bias will consciously or subconsciously restrict the uptake of some information from the world and privilege other information. Biases are cognitive devices.

The above argument assumes that natural selection evolves optimal solutions and also implies that bias X is an adaptation. Henrich and Boyd also make it clear that they are talking about genetic evolution with respect to these biases (1998), see also McElreath and Henrich (2007). However, Richerson and Boyd accept that the transmission may also be in part cultural (2005).

It is unclear what the source of the information coding for bias X is, and what are the proximate developmental mechanisms by which bias X emerges in the developing mind. The Henrich/Boyd/Richerson story is incomplete. Even if it can be shown that bias X is optimal in any environment whatsoever, and Henrich and Boyd have suggested that the conformity bias is favoured in a range of conditions broader, indeed, than those that favour a substantial reliance on social transmission (1998), then it still does not follow that natural selection for *genes* has shaped this bias. Selection could be acting upon a plastic cognitive device and upon culture.

What I will argue is that we can come to possess bias X without genetic natural selection for the X module adaptation. These biases could have cultural origins and result from social processes or human niche construction.

This is not a nature-nurture debate. No one denies that both genes and environments have profound influences upon the developmental outcomes of

human minds (Matt Ridley 2003, Dunbar and Barrett 2007, commentators to Quartz and Sejnowski 1997). I will argue that much of our mind is indeed developmentally plastic, and that many cognitive biases may be contingent on certain environments. I will outline the reasons I have for believing this. These reasons are in part theoretical constraints and in part empirical observations. There are good reasons to think that transmission biases cannot be genetic adaptations. This discussion of cognitive development will support the claim made in chapter two that the biases are not universal in space or time.

In order to determine the actual nature of these cognitive biases we can begin by examining the constraints that bind their development. I will note that there are tracking, neurobiological and learning theoretic constraints. I will describe the effects of each of these constraints before concluding that the mind cannot be a patchwork of genetically evolved modules. It is simply far from obvious that transmission biases are genetically encoded.

3.3
The tracking constraint

Richly specified innate modules cannot track rapidly changing targets. Richerson and Boyd argue convincingly (2005) that the Pleistocene was a period of unpredictable climate and, therefore, habitat change. Importantly, there would have been other environmental change, too, often as a result of human cultural behaviour. Mameli describes these changes (2007). They include significant changes to food composition and amount, social group size, mobility, accommodations and life rhythms. This change would have fed back in the form of selective pressures on humans and would have selected for early humans who adapt faster than genetic modules would have permitted (Sterelny 2003a, Laland 2007).

It seems that natural selection ought to select for plasticity of cognitive architecture in the face of changeable environments (especially the cultural) and the constraints of a genome with finite information restrictions (Sarnecki 2006, but see also Marcus (2004) for the view that the genome is not informationally impoverished).

Finally, Andy Clark (2003) has views of 'extended cognition', where our cognitive processes quite literally can be seen as looping out into the environment. We use pen and paper for extended memory, we rely on communication devices, we act in ways that are triggered by the external context, and we manipulate collections of ideas on screens and as notes on palm pilots. If this approach is a useful way of looking at the processes of mind, then the cognitive process is distributed across the environment. This environment is a rapidly changing target. This means that cognitive devices must be flexible and constructed in response to the environment. The changing nature of our cultural environment, particularly in recent times, coupled with the intimate relationship our minds have with our world, suggest that hardwiring all cognition genetically would be a bad evolutionary strategy. The need to track a changing environment must be a constraint on the nature of the transmission biases. If there is any innate specification it is likely to be innate under-specification of biases.

3.4
The neurobiological and learning theoretic constraints

Environmental processes will constrain the possible nature of cognitive biases, but so too will the nature of the brain itself. The acquisition and development of any bias must be consistent with what we know about neurobiology and learning. I will now detail the paradox that emerges when we try to reconcile the obvious lack of pre-wiring in the developing brain and the unreasonably slow pace of learning that a general purpose device ought to have. In order to

suggest a solution to this paradox I must first detail two opposing views on neurodevelopment, selectionism and constructivism. I will argue that taking a constructivist view permits us to explain the learning paradox with respect to cognitive biases without the need to fall back on unlikely genetic modules. I will also outline and reject Richard Samuels defense of nativism. I will reject cognitive arguments for the pre-wiring of psychological devices, in particular the frame problem. This will all suggest that cultural transmission biases are not genetic adaptations.

Learning is clearly important in the development of human minds. There are two issues here, learning at the level of cognition, and structural changes that occur at the level of neurons. If the human mind is the brain then these two processes must be intimately correlated. So much so, in fact, that no theory of minds can be complete without satisfying the *neurobiological condition*. The neurobiological condition is comprised of constraints upon mental function that are imposed by neural architecture. This may be seen as the fashion in which neurons in the brain are arranged and wire together in response to input according to our best empirical evidence. Any learning or developmental mechanisms postulated must reduce to neurobiologically plausible mechanisms.

This first section of my rejection of significant nativism relies on observations of brains and some theory about different methods of learning. There are two broad schools of thought regarding neural modification during developmental learning. These are *selectionism* and *constructivism*. Selectionism basically claims that there is an over proliferation of neural connections, predominantly during in utero development, and that these constitute the potential representation space of the mind. These constitute the learning resources. Connections are *eliminated* in a manner analogous to natural selection as a result of experience, and what emerges are the structures that underlie cognition. Hebbian learning theory, where neurons that 'fire together wire

together', underpins the persistence of certain neural connections. Constructivism, discussed below, postulates that the representational complexity of the brain (as evidenced by the quantity of neural connections) is much less rich to begin with and is built up by axonal and dendritic growth and the *formation* of new synapses over time[5,6].

Quartz and Sejnowksi (1997) argue in favour of constructivism. They detail two major features of the human brain that must be taken into consideration when discussing the emergence of our cognitive abilities. These features are as follows: (1) the developing cortex is largely free of domain-specific structure, and (2) that the representational properties of cortex are built by the problem domain confronting it.

The idea that the brain is altered by interaction with the environment is not new. But the traditional view was one of an abundant production of neurons early in development and the pruning of neural pathways that were not subsequently used. This was the selectionist model of brain development. Selectionism is consistent with other developmental processes in the body. The overproduction and attrition process is standard in developing neuromuscular synapses, in the neovascularization of body regions with new capillaries, and in the screening of T-cells in the thymus. However, presenting a wealth of neuroscientific data, Quartz and Sejnowksi suggest that the process of cortical change is *constructive* rather than selective and that these developmental

---

[5] It could be argued that the apparent gene shortage in human DNA is evidence that the wiring of the brain is not pre-coded. After all there are as few as 30,000 genes in the human genome. How could they possibly code for the trillions of connections. Gary Marcus takes up this issue (2004). He argues that genes act basically as a set of IF-THEN instructions. The complexity of the brain emerges thanks to contingent triggers on cascades of gene activation and deactivation. Marcus gives us a thoroughly detailed explanation of how the brain develops according to the genetic program, but this merely tells blank-slaters what they already assert: genes act in accordance with predictable environmental inputs (predominantly in utero) *in order to produce a slate* (which is blank). There is pre-wiring, Marcus admits, but not hard-wiring, the brain is still plastic. I will return to this pre-wiring later.

[6] Representational capacity is taken to co-vary with neuronal and axonal numbers in these models. This may, of course, be criticized as a tendentious assumption.

changes *affect the learning device itself* not just the representations contained in the brain. It is with these two points that these constructivist researchers diverge significantly from the preceding doctrine.

The traditional model of s*electionism* posits a haphazard oversupply of neuronal connections in the newborn according to a genetic program. These neurons are then trimmed into useful circuits by the agent's interaction with environmental stimuli. Selectionist development is, therefore, seen by constructivists as a reduction in potential representational capacity. Some selectionists do posit a cyclical overproduction and selection process in stages (e.g. Dehaene-Lambertz and Dehaene 1997) but importantly, according to most selectionists, the learning mechanisms remain unchanged by this process.

Constructivism on the other hand posits a *non-stationary learning device* which is sensitive to interaction with the environment, and also posits the sequential build up of *richer* representational structures over time. These more complex structures develop due, in part, to the large number of new *dendritic*[7] rather than axonal connections and are modulated by peripheral inputs. The peripheral inputs are signals arriving from sensory transducers such as visual, auditory, and proprioceptive organs among other peripheral stimuli. Some of this is further discussed below.

Learnability Problems

Arguments by Gold (1967) demonstrate that using a strategy of exhaustive hypothesis generation and then sequential trial of hypothesis allows a general learner to infer the grammars of a language. There is a limit, however, to the complexity of such hypotheses because a general learner soon strikes a version

---

[7] It is important to distinguish dendrites, which are small cell-body processes (of which a given neuron may possess a thousand or more) that receive input via other neurons, from axons, which are the longer, often myelin-covered output processes of neurons.

of the frame problem[8] when attempting to formulate the hypotheses to select from. Although results are theoretically inevitable, given enough time to try every hypothesis one by one, learning is *practically* impossible because of this vast search space. Hence, the argument goes, humans are *not* general purpose learners. We learn very complicated hypotheses about our world relatively rapidly. This entails a degree of *pre-specification* of knowledge.

Two factors can make learning difficult; these are a lack of bias and too much variance. *Bias* is the nearness of the representations that the learner already holds to the target of learning. If the learner is given a lot of important and relevant information before attempting to learn then learning will be easy. *Variance* is how far from the target representation a system's state lies. If there are too many possible learning solutions (only a few of which are good) then, without bias, the arrangement of the learning device is not likely to be close to any good target. A large network has a high variance, it is overwhelmed by possibilities. But on the other hand, highly biased networks will learn only in specific domains. These factors favour a more nativist, modular mind in which a cluster of small, biased networks each learn in a specific domain. This is the *learning-theoretic constraint*. According to this constraint unconscious cognitive biases ought to be innately pre-specified.

The Cortex is not Innately Specified

However, appealing to the neurobiological constraint, Quartz and Sejnowksi suggest that the physical cortex is no such device. Their arguments (1997) are backed up by more recent data on the cortices of preterm and neonatal babies (Kostovic and Jovanov-Milosevic 2006). I outline this recent data now.

---

[8] Traditionally the frame problem describes the difficulty of writing formulae that describe the effects of actions without having to write extensive formulae that define the mundane, obvious non-effects of those formulae (Shanahan 2004). The version of the frame problem considered here involves a learner who can systematically infer what the solutions to a problem might be, and then undertakes a sequential test of hypothesis over all possible solutions. Ultimately the learner will strike the optimal solution but perhaps not before significant time has elapsed.

Kostovic and Jovanov-Milosevic describe the development of the cortex in the following way. Following extensive growth phases in utero, there is retraction of many cortico-cortical axons (especially in the corpus callosum which connects the two cortical hemispheres). There is axonal arborization and growth of many short cortico-cortical connections. The thalamo-cortical tracts (those that convey input from the thalamus to the lowest layers of cortex) rapidly develop dendrites, spines, and synapses, beginning in the cortical layers that accept thalamic fibres and progressing in wave-like fashion to higher levels. *Sensory driven activity shapes these connections*. A rapid increase in dendritic synapses is seen in frontal cortex and continues for months (years?). There may even be gating of sensory input to primary cortical areas in order that intra-cortical circuitry elaborates in a controlled, gradual fashion. The empirical evidence continues to emphasize a lack of pre-wiring or its immediate wiring over. This offers no support for an innate modular approach to biases. Experience is important.

This is evidenced in the fact that the nervous systems of genetically identical grasshoppers, though similar, are not identical. Also, rats raised in complex environments have thicker, more connection rich cortices than those raised in bare cages (Klintsova and Greenough, 1999). Development definitely has a significant effect on cortical wiring and in *increasing* cortical complexity. This does not seem to fit with a purely selectionist picture of axonal trimming and programmed neural cell death, because according to strict selectionism cortical complexity decreases with time. I argue we have reason to doubt not only Evolutionary Psychology but also the selectionist picture.

Further to this data, transplant experiments where different regions of developing cortex are switched, rewiring cortex across different sensory modalities, and micro-anatomical studies suggest that cortical circuits are largely equipotent. Indeed, Mountcastle (1978) suggests that the entire cortex

operates according to a common algorithm and that regional variation is insignificant. Evidence for this view includes the existence of repetitive arrays of identical cortical columns and common circuitry. This data in favour of structural and operative uniformity of the cortex suggests that it is a substrate for constructive processes facilitated by dendritic and axonal connections, and that it does not house native databases or fixed pre-wiring. Afferent input is required for cortical development, there is no fixed regional specificity, and post-natal cortex is largely uniform. These facts comprise the neurobiological constraint.

3.5

Nativist responses fail to reject the neurobiological constraint

What we have then is a situation where, if there were genetic specification for modular adaptations, then the modules are not in place by the time that environmental influences are changing cortical structure. Module proponents will argue that the modules develop according to the environmental inputs (they are, after all, not genetic determinists) but then it seems that what has been pre-specified is cortical flexibility and this is not the same innate modularity thesis that early Evolutionary Psychologists proposed.

It could be further argued that equipotency doesn't mean cortical connections are not pre-specified. After all stem cells are equipotent and yet the result of their development into different body cell types is probably genetically coded. However, in the case of stem cells there are only one hundred or so possible body cell types and this number of outcomes could easily be pre-specified according to a simple set of rules.

There may well be *some* pre-specification of important cortical connections, but there is no doubt that 30,000 genes cannot pre-specify trillions of dendritic connections. The connectivity of the cortex is experience and context-

dependent. It responds to its environment just like the micro-vascularization of the heart responds to physical exercise. Only the grossest features of cortex are specified by genes, such as the fact that all cortices have six cell layers, and all visual cortices take inputs from the retina. This is specified genetically in the same way that all infants have four limbs and all hearts pump blood into the aorta.

Samuels (1998) discusses and rejects the neurobiological argument against representational nativism. Representational nativism asserts that some mental content is specified prior to learning. Samuels' argument could easily be generalised as an objection against architectural nativism of cognitive biases too. Architectural nativism asserts the existence of genetically determined structures and predispositions at a fine enough grain of analysis to account for cognitive devices like transmission biases. This would be 'hard-wiring' of the biases. First, Samuels claims that innately specified properties of neural tissue need not be intrinsic to the tissue. Second, the neurobiological constraint is too strong in its assumption about the inflexibility of innate structure.

The first point is saying that nativism may still be true even if the properties of the tissue are extrinsically or relationally specified. Just because cross-wiring of the visual cortex to the auditory cortex results in visual representations and the appearance of typical visual cortical structure in the auditory cortex, this doesn't mean that there is not innate specification. The specification may be spatial, or relational, or dependent on inputs. I accept this. But, if extrinsic properties are important, then different *cultural* environments could result in different cognitive architecture and therefore a range of biases to the flow of cultural information. Whether one still wants to call this 'nativism' or not is up for grabs, but either way extrinsic effects are relevant to architecture.

Samuels' second quibble with those who claim that there is no representational nativism is to offer a comparison with non-cortical tissue. If we transplant cells

that are innately specified to become teeth into the abdomen, then they become abdominal tissue. This, Samuels argues, shows that there is no 'principle of invariance' that is required for innate specification of human tissue because no one (he claims) will say that teeth and abdomens are not innately specified. But there is shifting in Samuels' example here. He is conflating the *tissue* that is specified to become teeth with the *region* that is specified to become abdomen. As all cells were originally of one type (pluripotent stem cells) then there is a point of no return in the developmental cascade beyond which certain tissues are set to become the organs that they actually become. If we were to transplant a tooth to the abdomen it does not become abdominal tissue, but if we transplant the cells that will ultimately become teeth to the abdomen then they do become abdominal tissue, thanks to their new location in the abdominal region. In the brain, particular regions may be 'innately' specified to become, say, the visual representation region merely because of the way peripheral nerves from the retinas enter the brain. Any brain tissue transplanted to this region will begin to receive inputs from the visual apparatus and can only then represent visual inputs. But this is uninteresting. There is obviously regional specialization in the human body. There is also obviously regional specification by genes in the human brain. Neurons from the retina always feed into the visual cortex for example. What is interesting is the fashion in which cells in a particular region perform their tasks and whether these mechanisms depend on heritable and learned cultural traits and contingent experiences, or are purely the result of genetic codes. Is the visual region a distributed representation of features in a visual scene or a photograph-like spatial representation? Or does it depend on how the individual developed? These are the interesting sorts of questions.

Overall, Samuels' arguments fail to convince that an anti-innate view of mental representations, let alone cognitive architecture is seriously flawed. I believe the neurobiological constraint holds and there is no reason to believe

that there is a rich innate representational structure in the brain. Context is not just significant, it is crucial.

But if we reject the nativist arguments then what we have is a paradox:

> The neurobiological constraint stands and suggests that much of the problem domain is not wired in ahead of time. The learning-theoretic constraint suggests that without bias we'd be geologically slow learners.

I will now spend the next few pages describing a resolution to this paradox.

3.6
Solving the learning paradox

Luckily, there is another way to learn that doesn't require minimizing variance or increasing pre-wired bias. An ordinary learner is endowed with a set amount of bias such that its initial state is nearer or further from the solution it seeks to learn. It possesses a set amount of variance such that the space of possible solutions is large or small. These configurations of bias and variance are the learner's *resources*. A *non-stationary learner*, on the other hand, is one in which the learning mechanism's resources are not time-invariant. Bias and variance can change over time according to experience. These constructivist learners build hypothesis space as they learn. Such non-stationary learners add resources (e.g. more neuron connections, or fresh arrays of the network) at a rate relative to their experience. This is an exponential property and slowly increases representational capacity by reducing bias at a rate that constrains variance. This makes the learner a *complete representation*, able to learn any learnable problem. Non-stationary learning is also a solution to the frame problem because the learning device can only represent a few simple solutions initially and as it learns simple representations it can expand its search space

accordingly. The simple representations learned initially, when variance is low, can bias the learning of more complex representations later. I will give some examples of overcoming the frame problem when I take this issue up under 'cognitive learning' below.

Quartz and Sejnowksi argue that post-natal cortical development carries on for far longer than is supposed. They infer that our cortical properties have evolved so as to 'maximize the capacity of environmental structure to shape [cortical] structure and function through constructive learning' (1997, pg 537).

It is interesting to note that Evolutionary Psychologists John Tooby and Leda Cosmides (2000, pg 72) accept that 'information and structure necessary for the proper development of an adaptation may be stored in the world as well as the genome'. Also, they describe the 'organizational domain' of a cognitive mechanism as those inputs that are required in order to organize the mechanism. The mechanisms seem likely to be elaborated according to data from the world. It seems to be a small step from Evolutionary Psychologists accepting this to their accepting of fully-fledged non-stationary learning and the possibility that biases are not wired in ahead of time. Our social environments could significantly influence cognition. Furthermore, social environments vary across cultures. I will provide evidence that this social engineering of cognition does indeed occur when I discuss environmental complexity and IQ below.

Cognitive Learning

Having looked at cortical elaboration and neuronal learning in development and argued that the nativist arguments at this level fail, I will now turn to learning at the cognitive level because there are two arguments in favour of the innateness of modular cognitive biases that relate to learning at the cognitive level. These are the *poverty of the stimulus* argument and the *frame problem*

(Sterelny 2003a). Sterelny discusses the poverty of the stimulus argument for several domains (language, folk biology and theory of mind) and concludes that many poverty of the stimulus arguments are weak (although they all have to be considered on a case by case basis). Important facts are, that we do not learn on our own (humans exist in an informationally rich social world in which generation N *teaches* generation N+1). Furthermore, generation N constructs an epistemically rich niche for generation N+1 to inhabit. Therefore, hypothesis space and representational relevance are engineered by our social worlds in order to enhance learning. Jesse Prinz (forthcoming) argues in a similar vein. Also, Dennett (2000) argues that we find words in our environment and these 'tools' allow us to develop non-language-based cognitive skills. The nativist modular response to these sort of arguments is to suggest that informationally engineered environments will help only once the right hypothesis is seized upon by the learner. The environment will help the learner to confirm the hypothesis. But again given the vast hypothesis space, this hypothesis needs to be pre-specified before it can be confirmed in this way.

I would reply to this on behalf of Sterelny, Prinz, Dennett, and others, by arguing that if it can be shown that formulating the right hypothesis without innate specification is not miraculous, then this objection to poverty of the stimulus arguments fails. If we can sensibly solve the frame problem we may have solved the poverty of the stimulus problem. Dwelling on poverty of the stimulus arguments is not going to help this approach. It is the frame problem that interests me more if proving that relevant hypotheses *can* be generated is the objective. And if relevant hypotheses can be generated for folk biology, language and theory of mind, then they are likely to be able to be generated for biases to navigate an environment of culture.

Recall that we are trying to explain the nature of psychological mechanisms that bias the uptake of information in the environment. In order for these

mechanisms to be learned by the brain, the brain must confine its search space or it will take forever to settle on a good solution. I will now show how a constructivist brain overcomes the frame problem at the neural level with an example from the motor skills domain. I will then move on to an example that uses a cognitive transmission bias, which emerges in response to social stimulation.

A constructivist brain would approach the frame problem by starting with a small array of simple, highly-biased, learning mechanisms. These first-order mechanisms must be genetically specified, but nothing *fancy* need be specified, if they are of small size with a narrow input domain this would suffice. These devices may be low-level cortex processing coarse sensory input. The low-level arrays each learn simple discriminations, indeed this is all they can do because they have significant bias on account of their small size and the nature of the sensory input they receive. Take a problem like a baby learning to direct the action of its arms. One simple learning device first learns to distinguish between arms-up and arms-down. A small neural network can learn such a simple discrimination with a minimum of equally simple visual or proprioceptive feedback. This neural network would form a first-order discriminatory device. Babies have poor visual acuity and it is features of their neurology such as this that suggest high bias in developmentally primitive devices. A parallel learning device may learn arms-left or arms-right. The outputs from these now biased (the output is binary) learning devices can be fed upwards to a freshly activated (added) representational device. There is no point *adding* (or constructing) this 'higher' cortical array until the inputs from below are consistent and coherent. But once added we see that the resources available to the learning device have changed over time. The new, second-order, device (though possessing no pre-wiring itself) is now biased by the inputs from below and can learn to discriminate oblique movements on the basis of vector addition between up and sideways. If the input 'up' is on, and the input 'left' is on, then the representation is 'diagonally up and left'. As

each higher level learns, further architecture is added and more complex tasks can be accomplished. Higher units may alter the gating of inputs to lower units so again the learning device possesses temporally variable resources.

An abstract example of the same process may clarify what is occurring. Imagine a problem in which the learner must find a six digit number but is only told after guessing all six digits whether they are right or wrong. Here bias is zero and variance is very large. Finding the solution is extremely difficult. However, if the learner can start by guessing only the first two digits, and is only given even numbers to select from, then the learning device for this first step of the problem is highly biased and variance is small. There are four options to choose from and only two occasions to choose on. This means that a systematic trial of sixteen options will find the solution. Only once the solution is hit upon do the next two blanks to be filled in appear. But now two out of four spaces are filled and the second step is no more difficult than the first step.

In the motor example patterns will begin to form in the inputs and desired outputs. For example to catch a ball the arms must be extended forward with the palms open. Once this association becomes clear then the two patterns can be coupled. The appearance of a ball in the air triggers the correct positioning of the arms. This is what happens when we habituate, learn manual skills, or form stereotypes. We can learn to associate proximate input cues that reliably lead to goal states and wire the behaviour to the proximate cue. This is the reward-prediction error-signal apparatus described by Montague (2006). We construct heuristics as we learn about our environments. It is likely that biases that shape the uptake of cultural information form in a similar way.

This mechanism of step-wise heuristic formation (where, for example, a child can learn motor patterns that are then activated by a proximal cue such as the appearance of a ball) is not limited to the motor domain. If a child first learns that directing attention at others is a good thing to do (because it is rewarded),

and then that following someone's attention to a target is useful. Then subsequently the child learns that giving someone a hug results in being told a story. And then learns that some stories are useful, then sees that other people give hugs only to certain individuals, and that these individuals are the reliable story-tellers. Then we have half of a prestige bias already. The child may tag the proximate cue 'someone gave a hug' to the action plan, 'pay attention to the recipient of the hug'. This transmission bias can be built sequentially and learning bias is provided by social circumstance. Social circumstance is inherited generation to generation but can change rapidly providing response to novelty. The learning process is channeled by social factors and constrained by events common in the niche. Eventually the bias becomes an unconscious habit. This is how we can learn without innate bias. Importantly, the biases created may vary somewhat from agent to agent. Such a mechanism also allows for environmental tracking as the device's parameters are not set for any particular environment ahead of time.

These heuristics approximate modules in that they are defined circuits that automatically activate when proximate 'buttons' (or information bearing cues in the world) are triggered. They are different to the Evolutionary Psychology conception of a module in that they are encoded in culture and the environment, and are malleable. By offloading information to the environment allowing it to provide bias, and running long developmental programs, human minds can be adaptively plastic. This enables adaptation to a large variety of niches. This would not be possible if our minds were composed of innately specified, automatic, encapsulated, domain-specific modules, because we would have to wait for genetic adaptation to catch up with need in order to adapt to novel environments.

It is not just skills that may be able to be built hierarchically like this, but language, concepts and heuristics. Once phonemes are mastered, words can be built. Once words are mastered sentences can be built. Once we can reliably

identify red, and wolves, then the story of Red Riding Hood can be understood. Constructive non-stationary learning is a powerful mechanism.

3.7

Predictions and empirical support for this position

I have just considered several theories of the development and nature of our cognitive mechanisms (we are of course particularly interested here in social transmission biases). It should be clear that there is a theoretical array of *possible* mechanisms by which the mind emerges from our evolutionary and environmental endowments.

1. Genetics code for cognitive phenotypes and these emerge in any environment (genetic determinism).
2. Genetics code for cognitive phenotypes and these emerge according to the environmental stimuli ('triggering' as in classical Evolutionary Psychology).
3. Genetics code for cognitive phenotypes and these can be overwritten by the environment and a plastic mind emerges.
4. Genetics code for plasticity and a cognitive phenotype emerges because of the current niche.

Hopefully my arguments have succeeded in suggesting that (4) is the serious contender. Of course, hybrid models are possible.

According to the present arguments, we should doubt that any given bias X is a genetic adaptation. There are competing hypotheses that explain the nature of the cognitive biases. I have argued that the following premises are true:

1. Human minds evolved in a highly non-stationary environment culturally, geographically, ecologically and climatically. This requires accurate tracking of the environment.

2. The *neurobiological constraint* suggests there is little bias hard-wired into developing cortex and that representational capacity is *constructed* in an environment-dependent way, not merely selected by attrition of neural connections.

3. The *learning theoretic* constraint favours a non-stationary learner in the human mind-brain.

4. That strict sense Evolutionary Psychology and strong cognitive nativism are false.

If true, these premises raise suspicion that the following will be found:

5. Cross-cultural differences in cognitive architecture.

6. Environmentally-dependent cognitive abilities (including external explanations of various biases).

7. Evidence of non-stationary learning in the brain.

8. Increasing difficulty in learning as the individual ages. This is because bias toward previously learned solutions increases, so the system finds it harder to reconfigure into new architectural forms without catastrophic rewiring. We also reach the anatomical limits of variance. This is where there are physical constraints upon the addition of new search space because the higher regions of naked cortex are already biased by previous learning.

I will begin the exploration of these issues by arguing that conditions (5) and (6) are satisfied in actuality.

(5) There are environmentally-dependent cognitive abilities

The intelligence quotient (IQ) measures the age-adjusted ability of individuals to perform certain mental tasks. It is established that the average population IQ has been rising over the last several decades since the tests were invented (Flynn 1987, 1999). In industrialised nations the rises have measured 1.33 standard deviations (SD) in one generation. In rural Kenya the rises between 1984 and 1998 were even more dramatic (Daley et al. 2003). Various hypotheses have been put forward as possible explanations of this. The following have all been argued as contributory: improved nutrition and health care, increases in environmental complexity, changes to family structure, improved schooling, and methodological factors. Data from the Daley study suggests that health in Kenyan children was worse in 1998 (though this is not the conclusion that these researchers draw from their data, it is undeniable that the incidence of significant anaemia, probably due to hookworm and treatment-resistant malaria, is evidenced in their study). Also these researchers have controlled for methodological issues unlike several of the western studies. For example the IQ tests given in 1984 and 1998 are unlike the Kenyan children's normal schooling. The same cannot be claimed in the west. The main remaining explanations are increases in parental literacy and environmental complexity.

It appears that, given certain environmental conditions, the phenotypes for higher IQs develop. These rises cannot be genetic adaptations because the timeframe is that of a single generation. These observations result from developmental environments interacting with genetics to form the final cognitive architecture of evolved adaptations (as per evolutionary psychology). However, the environments are presently optimal, for developing cognitive devices for high IQ, in ways that humans have not been exposed to in the past. For example, environmental complexity is a function of such modern

inventions as television and print media. Also the Kenyan populations are only recently literate.

The point here is that human niches significantly affect cognitive capacities. Synchronous and immediately successional human social groups can differ dramatically (up to 1.75 SD in continuous traits) in certain cognitive abilities. This is not consistent with highly-specialised, domain-specific, hard-wired, automatic, ancient evolutionary adaptations. At the very least this evidence demonstrates that niche engineering can significantly alter human cognitive traits. This is consistent with the neurobiological constraint discussed above.

(6) There exist cross-cultural cognitive differences

Adaptations resulting from selection over deep human evolutionary time ought to manifest as similar traits across human cultures. This, indeed, is some of Richerson and Boyd (2005) and Henrich and Gil-White's (2001) evidence for the conformity and prestige biases respectively as adaptations. But cross-cultural cognitive similarity has been called into question by a variety of recent data. Montague (2006) argues that Chinese players play trust games differently to North American players as evidenced by fMRI differences in simultaneous scanning during social interactions. Also, Pica et al. (2004), detail the results of numerical cognition experiments with Munduruku speakers in Brazil. The Munduruku language is numerically impoverished, with no discrete words for numbers greater than 5. Although they are able to successfully estimate the relative magnitudes of large numbers of items, Munduruku participants fail exact arithmetic with numbers greater than 4 or 5. This is an example of a cultural difference (language) causing a cognitive difference (arithmetic). There is a growing body of evidence that suggests there is social influence on the development of cognitive mechanisms. There isn't space to detail all the cross-cultural studies here, but see, for example, the variety of timeframes for acquisition of theory of mind (Vinden 1999), or a different *type* of theory of

mind (Lillard 1998), or the existence of languages without recursion (Everett 2005), or different strategies for simple arithmetic (Bjorklund, Hubertz and Reubens 2004), or memory performance (Carr et.al. 1989), or spatial cognition (Segall et al. 1966).

(7) The brain is a non-stationary learner
(8) We find it harder to learn as we age

These last two predictions await empirical confirmation. In particular (8) is anecdotally supported, but empirical data here will need to control for motivational confounding.

3.8
Maladaptivity, mechanism, and adaptation

The obvious questions generated by this approach to cognitive biases are the following:

1. If we're so plastic, why are we so often rigidly maladaptive?
2. *How* does environmental input shape cognitive devices? (at both the neural and cognitive levels)
3. What exactly is the adaptation here?

In answer to problem (1), maladaptation, at least Evolutionary Psychology had that one covered. If our cognitive biases evolved in the Pleistocene then current input conditions may vary such that modules are triggered by inputs not in their usual domain. If, however, neural constructivism is true, how can it be that we construct *maladaptive* biases? There are several possible answers to this question but broadly speaking the pace of cultural change can be so great that the structuring generation N does for generation N+1 can already be out of date. Similarly, it is possible that cultural ideas exist for their own benefit in

some sense at the expense of the host minds they help to shape (see 4.3). Also, the cognitive hardware need only be adaptive often enough to confer a selective advantage to the individual or group, it need not hit the right solution every time. Finally, no amount of plasticity will get perfect learning. There are always going to be cost-benefit trade-offs in the acquisition of cognitive devices.

Answers to (2) at the cognitive level are perhaps easier than the neural explanations. Take prestige bias, successful people are made obvious. We see them on TV and billboards, we are told that they are the best, competitions highlight their performances, they become salient and obvious. We can't help but copy them. They are prominent and the instructions to do as they do are pervasive. We can learn the proximate cues of deference such as, being on TV, having lots of people looking at them, receiving gifts and attention, as indicators of targets worth copying. We can use these cues as initiators of our copying programs. Once we have learned (in stepwise fashion) to copy prestigious people, then we can make the act automatic and unreflective, it becomes a fast and undemanding transmission bias. Conformity could be constructed in a similar fashion. We are told 'don't do that you'll stick out like a sore thumb', or 'no one does that'. We are punished for non-conformity (not wearing a uniform to school) and learn to value conformity over non-conformity. The bias may not be innate, rather the result of Pavlovian conditioning.

I've suggested that social environments cause the cognitive biases, but we must bear in mind that the cognitive biases also construct, in turn, social environments. We would expect to see that some biases are driven into the genome if they are consistently adaptive in any social environment that might be constructed. Are there any cognitive mechanisms that are always adaptive? Language, imitation, and folk physics may be examples. Perhaps also a general bias to be interested in the activities and beliefs of other agents may be

universally useful. However, in the last chapter I explained and gave examples of contexts in which the cultural transmission biases (e.g. conformity and prestige) will be maladaptive or expensive. These biases are not likely become genetically coded over vast expanses of time if human social environments keep changing. It is much more likely that cultures encode these biases not genomes.

In answer to (3), I suggest that natural selection has selected for an increasingly plastic cortex, which can learn in non-stationary fashion, and has an increasing capacity for variance. This variance is constrained by bias in the cultural niche. Combined with the apparatus of social information transfer this would have meant that culture was a tool for invading niches. Culture and cognition have certainly co-evolved. It is well accepted that cultural mechanisms drive the evolution of within-group similarities and between-group differences (Durham 1991). If culture encodes cognition then group selection would have driven the evolution of cognitive biases coded in culture. Indeed, if we construct certain features of our environments and social structures with enough reliability then genetic mechanisms may mutate out of existence as did vitamin C production in primates. If there *were* innate biases (perhaps in primitive brain regions) then they may have mutated away as the neocortex took on the role of constructing cultural-context dependent cognitive mechanisms that shape cultural flow in useful ways. Stephen Mithen (1996) has some interesting ideas along these lines and argues for an evolutionarily recent break-down of massive modularity as selection favoured more 'fluid' cognitive devices.

There is no reason to believe that the cognitive biases responsible for shaping the cultural flow of information are fixed. It seems theoretically possible (assuming the arguments against Evolutionary Psychology work) that we could engineer our environment (teach, mould, inform, provide the right input) such that our children develop the biases that we want them to. This opens the door

for social cognitive engineering. The human brain really may be a cultural artefact (Taylor 2006). As is the case with moral cognition, human cultural competencies develop reliably because they are stabilised by the cultural environment.

I have now explained the origins of human culture and the consequences of cultural transmission for the social organisation and psychology of our lineage. What remains now to complete an explanation of human culture is an analysis of what culture itself actually is. What is it that our biases are actually operating on? In the next chapter I put forward three possibilities, and in chapter five I defend the culture-as-memes hypothesis.

Chapter 4: Three Approaches to the Nature of Culture

Overview

In the last few chapters I've shown that there is a flow of socially acquired information among human agents, and that this flow provides adaptive benefits for people. I've suggested that biases must exist in order to solve the problem of individual adaptive choice when agents are confronted with informationally complex environments. I have also suggested that these biases are likely to result from the structure of cultural processes themselves. But what is this cultural information? How is it structured? What is it 'made of' so to speak?

There are three broadly evolutionary approaches to the nature of culture itself. I will now outline all three. I will begin with Dan Sperber's 'epidemiology of representations'. I will question his approach on the same grounds that I questioned Evolutionary Psychology in chapter three. I will then detail the Richerson and Boyd 'population dynamic' approach, but I find that they wrongly assume that cultural copying is of low fidelity. Finally, I will explain the memetic approach to culture and argue that the positive case for memes cannot be ignored. I will follow this exposition of the three evolutionary views of human culture with a thorough defense of memetics in the face of its major objections. I will conclude at the end of chapter five that memetics is a plausible model for the nature of culture. This is particularly likely when we consider that there are strong reasons to reject the Sperber and Richerson / Boyd approaches. Our psychological biases are operating on replicating cultural entities, or memes, that obey the algorithm of Darwinian natural selection.

4.1

Dan Sperber: An epidemiology of representations

Dan Sperber (1996) argues that culture is information stored in human minds and that the study of culture is that of the distribution of relevantly similar representations in the minds of human agents, and the public representations of that stored information. Sperber believes that information is transformed by human minds in every transmission event. Recall is not storage in reverse, and we don't copy cultural traits from other individuals with high fidelity. In fact, argues Sperber, we don't copy at all.

When culture is transmitted, the representation in one mind permits the creation of a public representation which then allows the creation of a new mental representation in the mind of another person. The public representation need not be a copy of the original mental representation and the new mental representation may be different again. This is the communication issue that Aunger (2000, 2007) and others raise repeatedly too. Sperber thinks it's fatal for cultural representation-*lineage* accounts of culture. Indeed, Sperber labours this point. In the transmission of culture, representations are altered almost *every time* they are passed on. There is nothing stable for natural selection to act upon, and not only that, there are no identifiable cultural trait ancestors.

Instead of copying, we *infer* the content of cultural information and these inferences are driven by evolved psychological attractors that privilege some types of representation over others. According to Sperber, human psychology fully explains human culture. There are no cultural traits that replicate or evolve, whether memetic units (e.g. Dawkins 1989) or non-discrete traits preserved at the population level (e.g. Richerson and Boyd 2005). What evolves, for Sperber, is human psychology. It should come as no surprise then that Dan Sperber favours Evolutionary Psychology as the correct approach to

human minds. At this point we must recall the many arguments against Evolutionary Psychology presented in the last chapter.

Any cultural replication that does occur is, for Sperber, just a limiting case of this extensive transformation. The brain is not a replication device, memory and communication transform information. An interpretation is a representation of a representation by virtue of some similarity of content, not because it is a copy.

Sperber argues for an *epidemiology of representations* as the correct model of culture. These representations have a similarity of content but are not copies. There is no replication and therefore no memes. However, as argued by Henrich and Boyd (2002), Gil-White (2005) and others, it seems replication may not be necessary for the evolution of culture if population processes compensate for error-prone transmission.

Ultimately Sperber compares two competing models of culture, the selection model and the attraction model (these could be called the evolving culture model and the evolving psychology model):

The Selection Model: In this model cultural traits replicate or copy and are differentially selected for according to fitness values or individuals' decisions. This is the model favoured by proponents of Darwinian approaches to culture in general, and memetics in particular. The problems with this approach, as Sperber sees them, are: a) representations don't replicate, they transform, and b) representations in one mind have too many 'cultural parents' for close analogies of genetic evolution to work. It is claimed that there is multiple source 'influence', ongoing mutation, and no lineage for selection to act upon. Temkin and Eldredge (2007) raise this issue too, and offer empirical data to support the multiple source influence phenomenon occurring in the temporal changes of material culture.

The Attraction Model: This is the model favoured by Sperber. According to this model, the outputs of a process of transmission are not wholly determined by the inputs or the choice of input weightings. If we put input A into a system, we get B out. B is not caused entirely by A but by the state of the particular system. Certain representations work better than others because variation or error is drawn back toward centres of attraction. Some representations are less cognitively demanding. For example we prefer to deal with whole numbers in economic transactions rather than fractions. 9.5% will be transformed to 10% for calculation purposes. There are 'transformation probabilities' that favour certain representations over others and this process is guided by our rich, evolved, innate psychology. Some cultural ideas have 'good form' e.g. Red Riding Hood (Sperber 1996, pg 108) and are more likely to have descendents. Factors that make for 'good form' are rooted in psychology and the cultural environment. I will present an idea similar to 'good form' when I introduce the *memetic index* in chapter five. The notion of the fit between representations and psychology seems important.

However, the attraction model seems at face value to be wrong because it would appear to have trouble accounting for cumulative aspects of culture. Attractors cannot change quickly enough to support rapidly accumulating culture. It is clear that culture accumulates, so how might this accumulating information be stored? The genome can certainly store information and preserve it across generations. But could information stored in the genome lead to such rapid cultural accumulation as we see? If genes coded for plasticity and there was a rapid change in environmental conditions then we might certainly see rapid behavioural *change* but not *accumulation*. It cannot be that culture is merely a manifestation of genetic evolution as culture accumulates far too rapidly for that to be the sole explanation. Similarly, if environmental variation caused culture to be evoked from minds then we ought to see similar cultures in similar environments. This is not supported by empirical data.

So, overall, Sperber argues for an epidemiological model of culture where those representations which are *cultural* are simply those individual representations, or representation similarities, that are more widely distributed. Cultural facts are distributions of representations and to say that various representations share a content feature is to say that they can all be interpreted, at a given level or from a given point of view, in the same way.

I will argue that Sperber is wrong, we do copy each other with high fidelity often enough for culture to accumulate and complex cultural adaptation to occur. The selection model is correct and both the replication and multiple cultural parent objections can be met.

Where is Sperber mistaken? He claims that, 'the information contained in the behaviours and cultural artefacts through which culture is transmitted *underdetermines* by itself the contents of the corresponding mental representation' (Sperber and Claidiere 2006). This is why he thinks all cultural transmission is constructive. Sperber feels that poverty-of-the-stimulus arguments (where learning something specific is impossible because of the paucity of accurate and relevant input) generalize to all forms of cultural learning. And he anticipates the reply of authors like Richerson and Boyd: Isn't the very existence of culture proof that there are mechanisms of information preservation effective enough to secure its relative stability? And Sperber replies: the convergence of constructive processes compensate for the limits of preservative processes. He thinks that rich psychological constraints underpin all constructive learning.

Sperber notes that 'learners can achieve what looks like strict reproduction when in fact the input information is incomplete and ambiguous, provided that their constructive abilities converge towards the same specific outcomes.' (Sperber and Claidiere 2006). Atran gives an example of this sort of process

(2001). If shown nine different sized circles on a screen then asked to copy them a group of people comes up with an array of different 'copies'. This is because the input information is ambiguous. It might be that it is a bunch of circles that is the important feature of the input, or the fact they are in a line, or they are poorly drawn bubbles. However, if told, 'these are the planets', then the copies the group makes are remarkably similar. This is because they now realise that it is important that there be nine circles and that the relative sizes are important. But this essential information is agent internal. It is not copied. The output is inferred. The agent-internal information could be propositional or it could be a cognitive mechanism. Sperber is emphasizing that rich psychology underpins all copying events and that such psychology is the target of natural selection in cultural evolutionary processes.

Sperber thinks that there are many cases where 'quasi-identical mental representations and quasi-identical behaviours propagate by causing each other in turns' (Sperber and Claidiere 2006), but claims that taking this as the paradigm of cultural transmission is a serious distortion. In this I think he is wrong, this is exactly what happens and I will endeavour to explain how when I argue in favour of culture-as-memes in chapter five.

4.2

Richerson and Boyd: Population dynamic models

Unlike Sperber, Boyd and Richerson (1985) and Richerson and Boyd (2005) accept that there *are* preservative copying processes at work in cultural transmission. Something is copied from one individual to another. This something is culture. But this copying is of low fidelity. This means that there is not *replication* in any particular copying event from any particular model, but population-level processes mean that information is retained with high fidelity at the level of the population. For example if a hunter copies the length of an arrowhead from his elders he may average their arrowhead lengths of six,

nine and fifteen centimeters, thereby crafting one that is ten centimeters long. He has not copied any one model with high fidelity, but the average of the information is retained in the population for the next generation to copy. The Richerson and Boyd view is that micro-level decisions (individual decisions of what to copy guided by biases) feed up to ecological scale effects (population distributions of cultural traits). Richerson and Boyd argue that imitation[9] occurs. When lots of imitation among a population is mixed with sufficient individual learning by trial and error then populations can adapt quickly. Replication of culture occurs at the macro-level.

According to Richerson and Boyd, culture must be encoded in material objects whether brains or artefacts. Culture is information stored in these forms. But without high fidelity replication of this information in individual copying events then bits of cultural information cannot be seen as evolutionary replicators. Therefore, Richerson and Boyd deny memetic theories of culture, which do assert that bits of cultural information are replicators.

Although they don't adhere to memetics as important in models of cultural evolution Richerson and Boyd elsewhere (Henrich, Boyd and Richerson, forthcoming) argue that (1) discrete replicator models of cultural inheritance can be useful even if mental representations are never discrete. And that (2) replicators are not necessary for cumulative adaptive cultural evolution. These assertions warrant some discussion.

(1) Henrich, Boyd and Richerson identify the objections that the cognitive anthropologists have made regarding memetics (real memes aren't discrete – Atran 2001, Bloch 2000) and those of Sperber, Atran, Boyer and others (the fidelity is too low) but argue that any *systematic transformations* of mental representations that occur due to cognitive attractors (as argued by Sperber and

---

[9] Richerson and Boyd use the term imitation generally to mean preservative copying strategies broadly construed.

outlined above) actually make discrete particle models of cultural evolution *better*. Henrich, Boyd and Richerson argue against the *existence* of memes, but they believe that models that treat cultural traits *as if* they were memes are insightful.

Henrich, Boyd and Richerson (forthcoming) describe examples of 'strong cognitive attractors'. These are features of minds which mean that the content of the representations held by those minds will tend to gravitate in a particular direction. For example given a continuous space of possibility from 'one hundred percent of the moon is made of green cheese', through to 'one hundred per cent of the moon is made of rock', it is more likely that any given agent holds a representation at either extreme of the continuum than it is that they hold a representation in the middle somewhere. This is because mid-way representations in this case are bizarre and incoherent (is the moon partly made of green cheese? Except for the rocks around craters?). Most people will hold cultural variants near such attractors. The attractors enable high fidelity transmission because noise in transmission will be drawn back to the attractor, and also provide the possibility of discrete-particle-type selection and evolution because of this predictable high fidelity transmission.

In Henrich, Richerson and Boyd's formal model, individuals who observe a cultural behaviour, described as a fractional number between 1 and 0, were modelled as tending to form representations nearer to either discrete representation 1 or 0. Additionally, selective attention was modelled so some representations were more attractive to the agent. Interestingly, this formal model assumes inferential transformation of the representation toward an attractor is much stronger than selective attention to a particular representation or representations. The result is that the dynamics of the system and the distribution of traits are closely approximated by discrete-trait replicator dynamics models.

Henrich, Boyd and Richerson conclude that we don't need to invoke memes in order to deploy replicator dynamics as an explanation of cumulative cultural adaptation, or to describe the nature of culture itself. And indeed, the stronger the inferential transformations the better the replicator dynamics approximation is. So the analysis is that mental representations (and culture) need not be discrete, yet discrete replicator dynamics are useful. I agree with this and will explain a pattern-based model of culture that is consistent with this argument.

(2) At first glance their second claim is obvious. If there is strict 100% replication then nothing can evolve. However, let's assume that there is merely error-prone copying. Henrich, Boyd and Richerson argue that even if a discrete trait has very inaccurate transmission on single case analysis, it may have high fidelity at a *population* level. If an agent copies another agent once, he may get it completely wrong. If he averages the behaviours of many agents then he is more likely to get it right. If we compare this crucial point to the upcoming discussion on memetics we will see that there is a definite distinction between this model and memetics. Tokens of cultural traits in population models don't have identifiable cultural parents of which they are copies; hence, there are no replicators as such if this is occurring. I will return to this issue of cultural parents in chapter five when I defend memetics.

Henrich, Boyd and Richerson also argue that fidelity of cultural transmission is generally low in any one instance of transmission between individuals (I will argue that it is high often enough) and that this tendency toward mutation may be represented by a value m in a dynamics model. The value m is a measure of the amount of error that has worked its way into the copied behaviour when an agent A copies the behaviour of an agent B. If A copies perfectly what B is doing then m = 0, however, if agent A fails to copy B in any relevant way at all then m > 0.5. The larger that m is, the less likely it is that cumulative adaptive evolution can occur. However, if the agent selects n different individuals to

learn from then this conformist transmission can correct for large errors. This is because each model is unlikely to be demonstrating the same errors (unless there is a cognitive or physiological attractor at work) so averaging the various behaviours permits the subtraction of error. Henrich, Boyd and Richerson conclude that fidelity of replication is not necessary for cumulative adaptation. They model this state of affairs with m = 0.2 in order to argue their case (Henrich, Boyd and Richerson forthcoming). They conclude that culture does not consist of replicating units or memes.

I think this inference is too hasty. The main shortcoming of the Richerson / Boyd picture is their belief that single-shot cultural replication is of low fidelity. Of course we need to ask what constitutes a replication event in order to argue about fidelity of replication, and this must be done in a principled way. If replication is merely copying once from one model then Henrich, Richerson and Boyd may have an argument for overall low fidelity of individual cultural transmission events. But if replication is a multi-factorial copying process involving several models, or one model on several occasions, then replication may still be occurring. It may be that an individual copies a model once and in this instance m = 0.2, the copy is 80% similar to the original. Observing the same model a second time allows the agent to refine the copy and it is now 85% similar. It may be that there is a sequential editing process at work. Compare the situation with genes. A replication event in the genome is not just the synthesis of a new strand of DNA. It is this, followed by (among other events) the checking of bases by exonuclease enzymes to ensure the copy is accurate. We can ask after an event if replication has occurred. We can also ask after several events if replication has yet occurred. Overall I agree, replication isn't required for fidelity (as Henrich, Richerson and Boyd have demonstrated), but I will argue that once we understand what a replication event is in cultural transmission then it turns out that transmission is of higher fidelity than Henrich, Boyd and Richerson are supposing. I will examine the

issue of replication in chapter five when I defend the high fidelity replication of memes. But first I will turn to describing the positive case for memetics.

4.3 Dawkins, Dennett and Blackmore: Memetics

Culture seems to evolve cumulatively by natural selection. That is the fundamental observation underlying the entire evolutionary approach to culture. Some theorists (e.g. Dawkins 1989, [1976]) claim that the cumulative adaptation by natural selection requires replicators. Genes are good replicators but, as Dawkins notes, good replicators need not be genes. Therefore if culture evolves by natural selection (and it certainly appears to) then there must be cultural replicators.

Traditionally investigators have looked for biological advantage in culture. Dawkins thinks that this approach is not necessarily fruitful. In following up the early work on cultural evolution by Popper, Cavalli-Sforza, Cloak and Cullen, Dawkins wrote that:

> The gene, the DNA molecule, happens to be the replicating entity that prevails on our own planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to become the basis for an evolutionary process…
>
> (Dawkins 1989, pg 192)

The gist of this is neatly summed up by Dennett (1995, after Lewontin 1980). It is important to appreciate that evolution by natural selection occurs whenever the following conditions exist:

(1) variation: there is a continuing abundance of different elements

(2) heredity or replication: the elements have the capacity to create copies or replicas of themselves

(3) differential 'fitness': the number of copies of an element that are
created in a given time varies, depending on interactions between the
features of that element and features of the environment in which it
persists

As Dennett has pointed out, this abstract definition of evolution by natural
selection is what Dawkins' memes exploit. The postulated memes are cultural
replicators that obey this algorithm and evolve in much the same way that
genes do. Dawkins again:

> Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making
> pots or of building arches. Just as genes propagate themselves in the gene pool by
> leaping from body to body via sperm or eggs, so memes propagate themselves in the
> meme pool by leaping from brain to brain via a process which, in the broad sense, can
> be called imitation.
>
> (Dawkins 1989, pg 192)

If memes are evolutionary replicators, then they *must* evolve by natural
selection. Additionally, there is no requirement that their evolution coincides
with human agent fitness. So it would seem that memes have a good chance of
explaining human maladaptive behaviour. In the same way that the lancet fluke
gets into the nervous system of an ant and drives it up a blade of grass to be
eaten by a cow in order that the fluke can reproduce in the cow's intestines,
memes could get into our minds and cause us to act in bizarre and biological
fitness reducing ways if such behaviour assists the replication of memes. The
question remains, are cultural traits replicators in a varying population and
with varying fitnesses?

Replicators

According to Dawkins, replicators are anything of which copies are made.
Good replicators are necessary for the evolution of complex adaptation and

good replicators exhibit high fidelity, fecundity and longevity. Replicators are a type and genes are a token of the type replicator. It is alleged that memes are another token of the replicator type. Importantly, two tokens of the same type need not share many features in common. For example, canned pineapple and canned tomatoes are both tokens of the type tinned food. But tomatoes and pineapples bear little actual similarity. If there are memes then they must satisfy the conditions for being replicators, they do not need to be similar to genes in every way.

Interestingly, Dawkins does not purport to be advancing a theory of human culture, merely arguing for the existence of replicators that are not genes, thereby universalizing his Darwinism. But perhaps this theory of memes can be pushed further. Could memes be explanatory of human culture rather than a feature of it? We must ask: how many memes are there? I will now present the positive case for memes.

Do memes exist? We can approach this question tangentially. In the biological sciences the existence of genes was inferred by hybrid experiments with pea plants. This coupled with the Darwinian theory of natural selection explains the phylogenetic lineages we observe through which we see the gradual (and punctuated) change of organisms' morphologies over time. The undisputed explanation of these phylogenetic changes are the effects of evolution by natural selection, and one of the targets of such selection are well characterised sequences of DNA which we can manipulate and, as a result, alter evolutionary trajectories. We know, therefore, that biological lineages (phylogenies) are based on genes.

In the cultural domain we see lineages too. I will briefly illustrate what I mean. Language changes by processes of descent with modification, just compare Shakespearian English with that of today. And it is not just change, linguists can trace the origins of words and their trajectories through languages and

cultures. We see also lineages in the technological domain. Temkin and Eldredge (2007) trace the phylogeny of cornet valves, cornets and psalteries. Lake (1998) examines lineages and the inheritance of village patterns and pots. Similarly look at the evolution of aircraft technology from the flimsy, clunky machines the Wright brothers flew, through to light, aerodynamic, pressurized passenger craft today. Ideas form lineages too. We can follow the physics of Aristotle to Newton and then Einstein. The fact that we see what looks like phylogenetics in the cultural domain is a reasonable indication that some evolutionary process is occurring (however, see Temkin and Eldredge (2007) for a discussion of why these lineages may not parallel biological phylogenies). It could be supposed, that evolutionary replicators are at work in the cultural domain too.

1. The cultural domain exhibits lineages of descent with modification just like the biological domain
2. Phylogenies in the biological domain are explained by natural selection acting on genes
3. Therefore, phylogenies in the cultural domain are explained by natural selection acting on a cultural replicator

Some Memes Exist

There are some unambiguous cases of memes. Even Dan Sperber, though he denies memes as a general explanation of culture, accepts that chain letters behave as evolutionary replicators. They are copied with high fidelity, they cause many copies of themselves, and being coded in material objects they have longevity. They can also evolve because they are not copied perfectly in every copying instance. Some chain letters survive because their threats or rewards are attractive, others don't appeal to our psychology and so die out, they are less fit. Dawkins (1989) argues that religion is effectively a chain letter. Religions consist of copying instructions and threats (or promises of

reward). The copying instructions are things like rituals to aid in memory of the scriptures, or instructions to spread the good word. The threats are those of, for example, hell fire. Dawkins argues that the resources we put into worshipping non-existent Gods for little or no benefit makes religion maladaptive for people. However, seeing religion as memes with their own fitness interests could explain important facts about human culture.

Here's a less obvious example of a meme. I may be shown a particular way of tying a knot by my father and then go off and use that knot to tie down a load on my trailer. In order for me to have done that then I must have formed a mental representation relevantly similar to my father's mental representation of the knot. I have replicated the mental representation. The 'knot' meme has been inherited. Also, this replicator can show variation. Perhaps I add a third twirl at the end of the knot in order to stabilise it better. If this works well I am more likely to employ the new representation in forming future knots. Perhaps I have less rope and I can only add a single twirl. My knot may end up a little different to my father's. The fitness of the knot representation *for the knot representation's sake* may also vary. Perhaps the single twirl knot is easier to copy so my son forms a representation of this simpler type of knot. This may happen even if it performs worse from my son's point of view. If given an array of possible knots to copy, my son copies (selects) the cognitively less demanding one. This simple knot representation becomes long-lived, spawns more copies of itself, and these copies are of high fidelity thanks to its simplicity. This cultural representation 'knot' has satisfied the conditions of the Darwinian algorithm and those of being a replicator.

Furthermore, on the selection and fitness of ideas, Kuhn has notably written, 'competition between segments of the scientific community is the only historical process that ever actually results in the rejection of one previously accepted theory and the adoption of another' (1962). Although Kuhn himself rejected the incremental model of scientific change, it is clear that there is

selection occurring in the cultural domain. That selection is actually pervasive. Dennett (1995) notes that we are bombarded by memes every day. We cannot pay attention to all the cultural information coming at us from newspapers, the internet, advertising, television and so forth. We necessarily select some culture over other culture. Not only are there obvious similarities between biology and culture, but instances of culture such as mental representations of knots and chain letters satisfy the conditions for evolutionary replicators. I argue that memetics is *prima facie* the correct way to look at culture. But many people disagree, so in the next chapter I will systematically respond to their objections. Taken with flaws in the 'epidemiology of representations' and 'population dynamic' approaches to culture, this will leave memes, as 'self-interested' evolving units, as the most likely explanation of human cultural evolution.

The question really is how many memes are there? The spectrum of explanation in memetics ranges from: there are almost no memes (Sperber), through there are many memes (Dawkins), to all culture is memes (Dennett, Blackmore).

Susan Blackmore (2001, pg 225) has written that, 'the meme is an evolutionary replicator, defined as information copied from person-to-person by imitation.' She builds her theory of memetics upon this foundation of imitation. According to Blackmore, if imitation occurs then memes exist. Any time imitation occurs a meme has been copied. Also, there is 'no requirement for memes to be symbolic or representational let alone conscious' (Blackmore 2005, pg 397). These broad definitions would include laughter and body-language-mirroring as memes. Such reflex-like human behaviours don't ordinarily seem to be counted as cultural. This is one area in which Blackmore's definitions are surely too broad.

Basically, Blackmore argues that once our ancestors became capable of imitation (in the specific sense of Tomasello 1999) there were two replicators, memes and genes. Memes are that which is imitated. But this definition of memes seems to miss the point that memes are supposed to be evolutionary replicators. This means that they must have fitness properties in order that they can evolve by natural selection. Blackmore is right to continue and discuss meme-gene co-evolution but her ideas about memes need to be bolstered by a discussion of memetic fitness. I agree with her that high fidelity copying exists, but this premise needs some argument. I will return to this in the next chapter when I argue that cultural transmission fidelity is often high and we are capable of single source copying.

Objections to Memetics

There are some important and often-raised objections to the view that memes explain culture (Aunger 2000, Sperber 2000a, Atran 2001). As I have noted, in order for memes to be a significant part of the correct explanation of cultural evolution then there must be identifiable meme lineages. If it cannot be shown that replication is occurring in the cultural domain, or the inheritance lines of memes cannot be identified, then there are no lineages for natural selection to act upon. I will reply to these replication and cultural lineage objections in the next chapter and show that the Dennett / Lewontin criterion (2), above, is satisfied.

It is also often argued that, even if there is replication, cultural transmission is of too low fidelity for particular memes to persist. Richerson and Boyd believe that the low fidelity of individual copying events is what necessitates the population processes that they model as underlying culture. I will argue in the next chapter that fidelity of copying in the cultural domain must be high and, indeed, actually is high.

The final crucial hurdle for memeticists is to demonstrate that memes have fitness values. Without identifiable fitnesses memes cannot be evolutionary replicators. If we are considering memes as possible targets of Darwinian selection then some measure of fitness value is essential. It is clear that some culture is fitter than other culture, for the human agent. For example, wearing fur is fitter than going naked in the Arctic. But some culture is fitter than other culture *for itself*. For example, American T shirts are fitter (spread more widely, have more appeal, and persist longer) in contemporary Africa than traditional tribal dress. We will need some measure of meme fitness that operates in a principled way. To satisfy this requirement I will present the idea of a *memetic index* to measure the fitness of memes.

A final objection to memetics is that it is explanatorily impotent. There is no phenomenon that memetics and only memetics can explain. There are other ways of ensuring similarity of learned behaviours across generations. I'll give three examples. (1) We may see such lineages if it were the case that agents capable only of asocial learning construct niches that cause agents in subsequent generations to learn the same behaviours that the original agents did. If one agent discovers how to fish for termites with a stick, his descendants are going to spend a lot of time around sticks and termite mounds. (2) Accurate replication at the population level can arise for reasons like averaging, blending or conforming. I introduced this theory in the previous section. In these cases there are no real lineages as there are no obvious 'cultural parents'. (3) It could be argued[10] that the evolution of a tool making *process* for example, or the creation of a reduction sequence, where a tool is sequentially modified, could have the incidental effect of producing what looks like a tool artefact lineage. Just because we find evidence that looks like a cultural phylogeny, for example an array of similar, but slightly modified stone

---

[10] Indeed, this has been argued by Ben Jeffares in, 'the Archaeology of Cultural Inheritance in Early Homo' a seminar at Victoria University of Wellington 22nd May 2007.

tools, this doesn't mean that there actually has been descent with modification rather than just modification.

In the face of competing explanations it looks like we need to say what memes *are* in order to prove that memes are at work as significant contributors to cultural evolution (rather than merely as curiosities).

There are many candidates for what memes are. They might be ideas, artefacts, behaviours, mental representations, or particular neural structures. Perhaps they are particularly bit-like, maybe as small as a word, a musical note, a mathematical symbol. But it does seem that we often learn larger, more cohesive chunks of culture from others, and that these chunks are inherited together in some way. A theory of memes will have to explain this. Are memes ideas, are they artefacts, or are they behaviours? I will argue in chapter six that all these things can be represented as *patterns* and it is these mental representations of patterns which evolve memetically.

Not All Behaviour is Memetic

It seems that there are at least some memes. But, it must be noted that even if memes exist then not all behaviour is memetic. It all *can* be memetic in that other agents can use our public representations to copy our mental representations, but we can also act without having copied anything from anyone in the same way that non-cultural animals can. There are instincts, innate behaviour, and there is non-social learning. For example, babies learn to walk and put things in their mouths to eat without the need for cultural transmission. We can also behave according to the results of individual trial and error. Similarly, goal emulation is not memetic, so not every instance of social learning involves the transmission of memes. However, as I will discuss, a lot of hybrid learning may still suffice for high fidelity memetic transmission.

If information is evaluated by an agent and altered to a great extent, then it is relying on existing knowledge and may produce a mental representation quite different from the mental representation that caused the behaviour. This is clearly *not* a case of memetic transmission because nothing is inherited. Social learning just means that memes can exist not that any given behaviour is caused by a meme. Something is a meme by virtue of its forming a lineage with greater or lesser success, with identifiable cultural parents, not because it is a behaviour, or because it is a mental representation, or pattern. I have put forward a positive case for memes, but it is incomplete. I will devote the entire next chapter to systematically replying to the critics of memetics.

4.4 Summary

In this chapter I have detailed three competing but perhaps not mutually exclusive views on the nature of human culture. We've seen that culture could be seen as a distribution of relevantly similar mental representations that result from the action of psychological transformers upon input data without there being copying. We've also seen that perhaps culture is best viewed as the sum of information distributed in a population, information that is indeed copied at the population level, rather than necessarily through individual copying events. And we've seen that culture may well be best seen as replicating units that seem at face value to obey the rules of evolution by natural selection.

These three views of what culture is best described as can be categorized according to how they each view the copying process in cultural transmission, and how they view the fidelity of copying.

	Copying	Fidelity
Epidemiology of Representations	None	N/A
Population Dynamic	Yes	Low
Memetics	Replication	High

In the next chapter I will meet the objections to memetic theories of culture and argue that most copying indeed occurs in replicative fashion and that the fidelity actually is high. The result will be that memes remain *plausible*. This coupled with the reasons for believing that Sperber is wrong because Evolutionary Psychology models in general are flawed, and reasons for believing that Richerson and Boyd misconstrue what a copying event is, will lead me to conclude that memes are also the *likely* explanation for the *majority* of culture.

In chapter six I will explain what memes actually are and how we copy them. Memes are mental representations of patterns in the world, and we employ a memetic stance in order to form such representations.

I will emphasize, however, that I do not believe that memetics, population dynamics and rich psychology are mutually exclusive. So, in chapter seven, I will conclude that the most useful model of culture is a synthetic one with important elements derived from all three. I will, however, also point out where each is clearly wrong.

Chapter Five: A Defence of Memetics

Overview

If memetics is to be a viable theory of culture then certain things *must* occur. Firstly, memes must form identifiable cultural lineages. This entails that memes must replicate and also that we are able to identify the 'cultural parents' of a meme. Secondly, the fidelity of replication must be shown to be high enough that the evolution of complex adaptation can occur. Finally, memes must have identifiable fitness properties in their environment. Memes must not only replicate but they must show variation and differential fitness. In this chapter I address these issues by showing that memes replicate, that we can often identify cultural parents, that the fidelity of copying is high, and that meme-fitness measures grant memetics explanatory power. I will introduce a memetic index of fitness features for memes; this will give predictive power to the meme theory. In the next chapter I will argue that memes are not only viable, but they account for lots of human culture. I will do this by arguing that memes are mental representations of real patterns in the world.

5.1
Objection 1: Where is the replication?

I will now demonstrate that memes replicate. If we side with Dawkins (1989, [1974]) then for evolution to occur there must be replicators. The question for memetics is whether there are replicators in the cultural domain. For Blackmore it is imitation that gives us cultural replicators. For Distin (2005), too, replication is the essential element in cultural evolution.

Criteria for replication are various but usually include a resemblance and a causality criterion (Godfrey-Smith 2000). Sperber (2000a), however, identifies *three* criteria for replication. These are:

1. B must be caused by A (together with background conditions)

2. B must be similar in relevant respects to A

3. The process that generates B must obtain the information that makes B similar to A from A.

It is all very well for copies to be made, but a *good* replicator, according to Dawkins, has the following properties: longevity, fecundity and fidelity. If these properties exist then they are sufficient for something to act as an evolutionary replicator, and therefore to be the target of natural selection.

When deciding what counts as a good evolutionary replicator we must be careful not to raise the bar too high. If we are too strict about what counts as a replicator then genes (the paradigm replicators) fail. Genes, it turns out, are not particularly 'gene-like'. Rather than being composed merely of discrete replicating entities, the genome is a hodge-podge of discontinuous, overlapping and interchangeable 'units'. We see introns of non-coding sequences. We see sequences where the same stretch of DNA partially codes for more than one protein. Some genes move around the genome, others are completely nested inside still other genes. We see alternative versions of exons spliced into a coding region. Indeed, as Mesoudi et al. (2006, pg 343) have noted, 'the modern concept of the gene is hence characterized as abstract, general, and open, with fuzzy boundaries.' If anything, genes are particularly culture-like! One is almost motivated to ask: are there genes? Of course this is absurd. There are genes and we can manipulate them. We must, therefore, note that the criteria for replicators are looser than those of the traditional conception of a discrete gene.

Dan Sperber (2000a) believes that replication is *essential* to an evolutionary process and argues against meme theories on these grounds. To illustrate his point he gives us an example of cultural transmission. He asks us to imagine a

random squiggle being copied from memory down a chain by ten people. He argues that the noise that would certainly work its way into the pattern after several 'replications' means that we could identify the order that these diagrams were made in, after they were drawn. He then argues that given a five-pointed star that the same exercise would yield ten similar diagrams that cannot be put into order again. The noisy transmission occurring with the squiggle, he argues, is evidence that replication is not occurring. The implication is that if there is not replication, if the copying is too noisy, then the thing being replicated is not stable enough over time to undergo cumulative selection (see 'the fidelity issue' below for the importance of this in evolutionary adaptation).

Sperber (1996, 2000a) denies that replication is occurring in the cultural domain. It is not obvious that we copy cultural units from others in a way that they can be said to 'replicate'. Is my paraphrasing of a folk tale's plot really replication? Are these then one or two memes? And if someone else were to paraphrase my version would it bear any similarity to the original? If I make pumpkin soup and then someone tastes it and decides to make their own version, is it still my meme for pumpkin soup? Or is there too much blending, averaging, creative recombination and information loss in cultural transmission so that it is not appropriate to talk of replication?

Sperber's squiggle fails his replication criteria on (2), it is not similar after each replication. And his star fails on (3), the information comes from our minds not the star. Sperber is betting that all cultural transmission events fail in one of these two fashions.

Here's the point Sperber is trying to get at. If you already have the information in your mind then you are not replicating anything, you are instantiating it. In his example we are copying the squiggle (poorly) and instantiating a pre-held star meme. But, surely no infants are born with the star meme already in their

minds. There was a time at which 'five-pointed star' established itself in any one person's mind. Sperber's example is comparing a novel meme with an already possessed one. The already possessed meme is a good meme (is fit in the environment of human psychology) by virtue of the fact that we all hold it and can call it to mind at will. The squiggle is a poor meme (less fit) and therefore is not copied easily. The persistence of the star meme in the population, as a symbol, is testimony to its fitness as a meme. Just because we infer structure (in the case of recognising the star) rather than copying, and just because some memes' reproduction is triggered from memory, does not mean that these mental objects are not long-lived, fecund and of high fidelity. That is, lineages still exist. Selection acts upon lineages. But note that if the meme is successful in its own right *and* useful to the agent then selection acts *not only* to favour the meme, *but also* to favour the mental mechanisms that tend to acquire that meme. Stars are selected for if they are good memes *and* mental mechanisms that acquire symbols like stars are selected for if they are good for the agent. Sperber focuses only on the second of these processes. Here is another example. We could try the same experiment with the story of Little Red Riding Hood. If we told people to copy the story as we began telling it, then long before it was finished the audience would recognize it and complete the tale. They already possess the Red Riding Hood meme and can instantiate it without copying. There would be causality and similarity, but no inheritance. A completely novel story on the other hand must be listened to entirely and then whether or not it is accurately copied will depend on such features as the attention it held, was it interesting or mundane, was it relevant to the audience, and so forth (see *memetic index* below). The squiggle, sadly, is neither relevant nor interesting, and its function is the same whether it has two spirals and a curl or three curls and a loop. It is uninteresting and therefore has a low fitness in the environment of human psychology.

So there are reasons to doubt the conclusion that Sperber draws from his 'objection to the memetic approach to culture'. We can argue in one of three

ways against his denial of replication. First, we can look at the star as an example of a pre-held meme. Such learned memes could be the basis of what appear to be cognitive attractors. But such attractors result from experience with earlier memes, not only from evolved psychological devices. Secondly, we can argue that replication in the case of the squiggle fails because it is simply not a good meme. Replication may only occur in the cultural domain when the target of copying is copyable. Finally, we can attack Sperber's criteria for replication and claim that it fails to pick out all cases of replication. Here is an example of replication that denies Sperber's condition (3).

When analyzing replication, Godfrey-Smith (2000) gives the example of *reverse translation*. This is the hypothetical process converting primary protein structure into nucleic acid sequence. This is the reverse of what occurs in normal cells. Ordinarily enzymes catalyse the translation of nucleic acid sequence into a sequence of amino acids that fold to form a protein. If any organism were to exhibit reverse translation then a hypothetical enzyme, *reverse translatase*, would convert protein to DNA. Indeed a cell possessing both mechanisms could convert DNA(1) into protein, then protein back into DNA(2). There is no doubt that the genes coding for the protein are replicators (this is part of the central dogma of biochemisty). But, ordinarily, DNA is replicated through a process of DNA transcription that generates replicates directly from templates of DNA. In the hypothetical process we seem to have DNA replication occurring via a protein intermediary. However, the DNA(2) strand produced after reverse translation is still a copy (within the limits of code redundancy) of the original DNA(1) sequence.

> What are the replicators here? The [proteins / phenotypes] produced are not replicates of the [DNA / genotypes], nor vice versa. This is because they are not *similar*. But is a stretch of DNA produced at the end of the life cycle, via reverse translation, a replicate of the corresponding DNA that started the life cycle? The answer is yes.
>
> (Godfrey-Smith 2000, my emphasis).

Similarity here for Godfrey-Smith is actual, physical, or substrate similarity. He is noting that the DNA and the protein are not physically similar. But, it could be argued that the primary protein sequence and the DNA bear an *information similarity* to each other as two encodings of the same information, like a song recorded as a musical score and a CD recording of the same music being played both encode the individual notes; one in symbolic physical form, the other in an auditory recording physical form.

It may be the case that mental and public representations of cultural traits, though physically different, share a sufficient informational similarity with each other. It is possible that, using the right mechanism, one could extract memetic information from the public representation (or expression of a meme) in analogous ways to how reverse-translatase (the imaginary enzyme) would extract DNA information from the protein, or someone learning a song by ear would extract the score from the CD. I argue below that our ability to employ a memetic stance gives human minds this capability.

There are two possibilities for what has been replicated in the above example. Either DNA has been replicated via a protein intermediate, or information has replicated at each step. In the second case the protein is an informational replicate of the original DNA(1) sequence and the second DNA(2) sequence is an informational replicate of the protein. If amino acids and DNA *are* (informationally) similar, then replication *can* be said to have occurred from *gene to protein*. If amino acids and DNA are *not* 'similar' then replication has only occurred from DNA(1) to DNA(2) without similarity of the intermediate substrate, and *is* occurring from mental representation to mental representation even if there is a public representation intermediate.

1. If there is *not* physical similarity then there is 'skipping' replication from DNA(1) to protein(1) to DNA(2) to protein(2) to DNA(3).

2. If there is informational similarity, then we have non-skipping replication.

Either way the DNA(n) lineage persists.

I hope I have illustrated above how even sequences of gene-protein-gene should be seen as copying, or 'skipping' replication as the information in gene(2) is drawn at least indirectly from gene(1), and gene(2) codes for exactly the same protein as gene(1).

Another way of approaching this problem is to remember Lewontin's criteria for evolution by natural selection (Lewontin 1980). Lewontin claims that the principle of differential fitness may be formulated as follows: Different variants leave different numbers of offspring either in immediate *or remote* generations (my emphasis). Provided that a subsequent mental representation is reliably and sufficiently similar to the original representation and is the result of a causal pathway traceable to the original representation then the meme for that trait can be said to have left descendents in a remote generation. If I make pumpkin soup and Bob copies me but gets it wrong, then nothing is reproduced. But if Anne copies Bob (gets the coarse-grained template of the meme from Bob) and her soup turns out the same as my original, thanks to the shared background Anne and I have and our reliably similar individual learning circumstances, then my meme for pumpkin soup has left a descendent in a remote generation. This sort of hybrid transmission can be responsible for persisting high fidelity transmission of cultural objects.

Overall, Sperber's replication criteria (3) is not necessary. There can be replication when the replicate doesn't directly obtain the information making it similar to the template from the template.

5.2

Objection 2: Where are the cultural parents?

This is a serious but empirically testable objection to memetics. We need to ask what is the cultural parent or parents of a meme. If there are too many parents and culture blends too much then offspring aren't going to resemble any particular parents. There will not be predictable or systematic relationships between parents and offspring. There will be too much blending and phylogenies will disappear as Temkin and Eldredge have argued can be seen in psaltery and cornet valve lineages (2007). However, it is far from clear whether the usual case of cultural transmission is from one or a couple of cultural 'parents' to one offspring, or from many to one. It is not clear whether children tend to learn from one source or several.

For example, most people's story of Red Riding Hood is not a copy of a *particular* version of the tale. We seem to pick up bits of the story from here and there, some from our parents, some from school, some from grandma (when she hasn't been eaten by a wolf). However, *even* if this is the case it is still not clear that there is not a discernable inheritance lineage.

It could be that the cultural parent is the instance of a meme that caused the agent to first possess the template for that meme. A template is simply a rough mental representation of the salient features of a cultural trait. This may be something (not quite) as simple as 'the story of a wolf who eats grandma and an axe-man who saves Red Riding Hood'. This template could be acquired from a single cultural parent in a single interaction. The acquisition may involve a single cultural parent, or a couple, or a few if 'cultural sex' occurs and we deduce some abstracted ideal template from multiple parents. The template is first obtained and possession of this template permits fast and accurate copying of any particular further instances of the meme. Once you understand the basic story of Red Riding Hood, you can easily copy a new

version in which the axe is switched for a baseball bat. If that subsequent version appeals to you more then your axe meme has 'died' and the baseball bat version has spawned another offspring. The base-ball bat version you heard will now be the cultural parent of your new meme. Memes will be in competition with each other for survival.

In the case of Sperber's star, a child will at some point form the 'star' pattern in their mind *for the first time*. They then possess the star meme as a pattern. They can appreciate other stars as instances of this pattern. When they come to produce the star themselves, they do so with the noise of performance so that others will recognize the new star as star + x % noise.

If it is argued that there are many cultural parents and averaging or blending occurs then I would argue, as Distin does (2005), that the child must first possess the meme in order to average or blend something with it. Information must have already been copied if averaging is to occur. If the averaging process changes the meme too much then the lineage is lost. If it refines it then the lineage persists. Even if blending means that some lineages become hopelessly corrupted, then this still doesn't mean that there aren't some memes that are passed on reliably, with high fidelity, from identifiable parents in our cultural transmission. This is likely to occur with vertical transmission where there are sole models for particular copying events. At best, the cultural parents objection weakens the scope of memetic explanation. What is really needed is more empirical data on the nature of our cultural transmission. Do we tend to copy from one source among many or from many sources at once?

5.3
Objection 3: Cultural copying is too low fidelity

The complaint that many raise in the face of memetic theories of culture is that even if there is replication, it is so error-prone that there could never be stable

replicator lineages. We must first see why this is an important concern, so I will begin with some theory about copying fidelity in evolution. Then I must diffuse the objection by arguing that there are at least four reasons to believe that copy fidelity is actually high. These rely on a 'correcting' process we employ, on world models that individuals share, on the learning environment and its structured nature, and the existence of language. If copy fidelity is high then it should be clear that the Richerson / Boyd models of low fidelity transmission with population-level compensation are superfluous or of peripheral importance. Also it should be clear that the Sperber view is wrong because something actually is copied between minds. These things are memes.

Dan Sperber feels that fidelity is crucial, and complains that, 'unlike genes ideas rarely copy with anything close to absolute fidelity' (2000a). But why does copy fidelity matter? If copying fidelity is low, the reasoning runs, then cumulative evolution allegedly cannot occur, therefore, there can be no complex adaptation of memes. Let's see how this argument is supposed to run.

Mark Ridley (2000) describes an analogy between the game of Chinese whispers and replicator dynamics. Because of the accumulation of mutations over time, if an average parent (either a whisperer in the game, or a replicator like a gene) fails to leave at least one offspring without a copying mistake then the replicator (whispered message or gene) mutates to randomness over time. Importantly, however, evolution and complex adaptation *can* occur if many offspring are produced only a few of which are error-free.

There is another theory that becomes relevant here. The error-catastrophe argument (Eigen and Schuster 1977) shows that given error-prone replication of a template in which the template then competes with the error-ridden copies of itself (even if they have a slower replication rate) then, over time, and once a threshold rate of mutation is reached, the template ends up no more common than any single variant. Basically, the replicator slowly loses self-identity

because of the increasing number of mutations in each subsequent copy. Error-catastrophe has occurred and the original template lineage ceases to dominate, or even goes extinct. Applied to memetics this suggests that if error rates are high then memes *effectively* disappear over time. If there is a mistake in each retelling of Red Riding Hood then after several replications it is no longer the same story, and is now competing with many different stories that were originally spawned from Red Riding Hood. However, error-catastrophe arguments need to account for selection. It may be that a single or a few faithful offspring is all that is needed to sustain a lineage, but only if the error-ridden copies are eliminated from the population. High fidelity *or* strong selection against mutants seems to be required. If the error-ridden versions of Red Riding Hood are not appealing (or in the retelling of a joke the punch-line is messed up) then there will be selection against these copies. The original (self-identity possessing) copy will persist.

Therefore, for complex adaptation to occur, mutations must be small because large mutations disrupt the replicator. The chance of a large mutation yielding an adaptive benefit is small. In the case of organisms, large mutations are usually fatal. In the case of memes, a large error or random change in the pattern ought not to produce a functional behaviour or artefact, or at least the resulting product is unlikely to be an improvement over its parent. Improvements are important because if each intermediate product is not fitter than its predecessor then adaptation will not occur. If mutations are necessarily small then cumulative selection is the only process by which adaptation can occur.

What of the case for memes? Sperber clearly thinks that the small errors that creep in during each replication are analogous to Chinese whispers where every transmission involves an error. But a person demonstrating a meme can be copied by several or very many others. It only takes one or a few such agents to copy it correctly for the lineage to persist. Even if most of the time

when we 'copy' culture we introduce error, there still may be enough faithful replications for adaptation to occur over time. This is because there then ought to be strong selection for the most fit, most adapted memes, and selection against the error-ridden copies. Also, even if expression of the trait is noisy, this doesn't stop the copier from identifying the underlying pattern or instruction set. This is just like the situation where we can see a pattern in a noisy pixilated image even if there are only a few rough dots to indicate what is supposed to be there.

Ridley makes a further important point. He notes that having two copies of each gene has protected the human genome from the accumulation of error. As it is unlikely that the same error occurs in each copy, the good copy can mask the bad. In cultural transmission we can observe (and copy) a range of different cultural objects. Even if there is frequently error we will come to possess some error-free copies of the instructions for cultural traits. By trying out different behaviours an agent can settle on the most successful. They will then tend to exhibit this meme as a public representation and this is statistically more likely to be observed and copied by others. The 'good' cultural trait can mask error-ridden ones. Even within our minds we may be trying and selecting memes in decoupled fashion. In seeing public cultural representations subconscious processes may be trying to determine the mental representation that led to this public expression. Many possible underlying instruction sets could be tested and projected into different situations and many discarded with only a very few ever reaching consciousness. If we make copying errors in identifying memes then these errors may be weeded out by comparing multiple versions of what the meme might be as more information is assessed. If multiple-draft type models of mind are correct then many copies of memes may never reach consciousness, let alone expression. This could be an error minimizing mechanism.

I have suggested that fidelity of the initial copying event could be quite low and we still could have evolution by natural selection if there is ultimately a reliable correspondence between parent and offspring, caused *by whatever means*. This surely could include individual learning that makes use of processes that are reliably similar between individuals. Similarly, Richerson and Boyd have shown that averaging during cultural transmission can enhance low fidelity copying. They model a 20% error value per copying event (see chapter four). I suggest that even without averaging processes, if patterns are copied, then even a high percentage of error may be able to be restored to the original pattern in the right circumstances. For example, figure 6.1 (next chapter) shows a pattern 'bar-code', which can be discerned easily even when 33% of the image is noise. We can see this expression and yet mentally represent the pattern 'bar-code' without the noise. If memes are pattern representations (as I argue in the next chapter) then we may perform this correcting procedure with cultural representations.

It is important to note that such pattern 'correcting' processes, where bar-code + noise can be restored to pure bar-code, do not mean that we are failing to meet the replicator condition. It could be argued that bar-code, copied to bar-code plus 30% noise, copied to bar-code again by noise elimination, is not replication because we have a sequence of $X - X^* - X$. And mutations (e.g. X^*) are not faithfully transmitted. Faithful transmission of new variants would be required for evolution to occur. But we must look at this at the right grain of analysis. Let's say there is no fitness difference between five black squares in a row with 25% noise or with 33% noise. But there *is* a fitness difference if we have four black squares plus a triangle. A mutation, therefore, occurs in bar-code if it ceases to be a sequence of five black squares. If it is incorrectly copied as four black squares and a black triangle, then we have a novel meme (and a new pattern), which may or may not be fitter than its predecessor. Think of pumpkin soup with one onion, versus pumpkin soup with two onions. There

may be no fitness difference in this case. It is just noisy expression. But pumpkin soup with carrots is a different meme.

I have just argued that correcting processes enhance fidelity of cultural transmission. The next reason to think that cultural transmission is of high fidelity is the existence of similar internal world-models that agents share. Such agent-internal world-models affect transmission. Take the following case. Two people A and B split up to head across town and meet at the airport. One must collect groceries on the way, the other the kids from school. Both are familiar with the city, they share the same internal representation. At some point A can ring B on their cellular phone and say, 'I'm stuck in traffic on Victoria Street.' Using their knowledge of the city (their internal model) and traffic patterns, person B can infer that A will be about 10min late. A short burst of information transmits a great deal of accurate knowledge if internal models are shared (see Montague 2006). On the other hand, if models are not shared much more information about the situation must be given. The transmission is more difficult. In the case of cultural transmission, if agents share a world model then fidelity is high. Often we have situations where members of generation 1 have a different understanding of the world to members of generation 2. If the members of G2 all share a model then horizontal transmission fidelity can be high with relatively little transmission cost. The point is that if environments change significantly from G1 to G2 then vertical transmission can become error-prone. But transmission *within* G2 can be enhanced by vernacular and fashion similarities.

The structured nature of the learning environment is a third mechanism that adds to fidelity. For example, complex skills are not copied all at once in a single transmission event. Skills such as canoe making, practicing medicine, or any sort of master-apprentice transmission occurs (usually vertically or obliquely) over time. The copying event for such culture must be seen as *distributed through time*. By copying first a coarse-grained pattern of the

overall idea, then copying one at a time smaller sub-patterns of each individual skill, with each learning step being constrained by the niche, available materials, shared world-models, and a constructed trial-and-error environment, accurate high fidelity copying of the representation occurs. This is the only way that complex patterns can be faithfully copied. Small errors are constantly corrected across time. At the end of the entire distributed copying event a mental representation has replicated. The representation exists now in two minds where before it was only in one. Of course, simpler skills and traits could be copied in a single step horizontally.

Finally, language may help to improve the fidelity of pattern transmission by highlighting particularly salient aspects of the overall pattern. For example I may see Santa Claus at a shopping mall as a child. Now the concept 'Santa Claus' is a culturally transmitted pattern constituted by the notions of a man, presents, Christmas time, a jolly laugh and so on. But what makes this pattern different from my father? If I ask, 'who's that mummy?' I might get the response, 'Santa Claus, he's the man in the red hat and suit'. The salient aspects of Santa Claus are highlighted and I can draw on the concepts 'red', 'hat' and 'suit' which I have encoded already.

As well as noting how fidelity is enhanced in any instance of transmission we must ask what constitutes a copying event. Different answers to this question will affect the answer to the question of whether cultural transmission is of high or low fidelity. If a copying event is merely a one-off attempt to understand the instruction set underlying someone's behaviour then there may be an argument for low fidelity of complex transmission. But if copying events include putting a meme into practice and comparing the resultant behaviour with the original modelled behaviour, again and again if necessary, until it is a match then fidelity is going to be higher. Just like DNA is edited by exonucleases, which check newly inserted bases for correctness, so memes can be edited by the agent. This is hybrid learning and it is as crucial to cultural

replication as exonuclease activity is to DNA replication. We fill in the blanks and excise the errors, just as the machinery of DNA replication does. Genes do not replicate *on their own*, and we should not require memes to either. Once an agent has inferred the program he can run the program and realise the behaviour. Because of our ability to imitate motor patterns (practised when we were infants and unconsciously enlisted as body language mirroring) we can compare our resulting motor pattern, our interpretation of running the program, with that of the target and see if we have it right. This is meme editing. Copying events must be seen as distributed through time.

Fidelity doesn't have to be perfect on the first go for the agent to eventually pick up a copy of the original behaviour. Indeed, in trying to copy Roger Federer's serve I might note that even he doesn't perform exactly the same motor sequence every single time. So I must be trying to copy some kind of abstracted ideal serve (see Gil-White 2005). This ideal may be a maxima or the mean of a statistical cloud of behaviour. Sterelny gives the example of his father's and his own understanding of the cricketer Bradman (Sterelny 2006a). The point here is that his own knowledge of Bradman was collected over many interactions with his father. We can approach this sort of situation in three ways. Either we can argue that the entire childhood-long interaction was the copying event for the Bradman complex. Or we can argue that multiple memes have been copied in multiple copying events. Or we could claim that a low resolution representation has been copied initially (e.g. Bradman, cricketer) and this has been filled out over time with new 'pixels' of information. These new pixels may have been generated by individual learning or have arisen from other social sources, such as old cricket almanacs, news stories, or radio replays. Either way, the hybrid of social transmission of a coarse grained representation from father to son, and then subsequent refining of the representation, result in a copy of the original cultural object. Also, it must be seen that the sequence of learning (from coarse grain to fine) is organized and optimized by one generation for the next.

Michael Tomasello (1999) suggests that imitation is of high fidelity and that this explains our cultural capacities. Hopefully it is clear now that imitation need not actually be all that accurate in order for memes to be copied accurately because of all the peripheral machinery of copy correcting, shared world-models, the structure of the learning environment, language, and the distributed nature of the copying event. Such things *do* make copying high fidelity. Richerson and Boyd (2005) accept the arguments for low fidelity copying put forward by the likes of Sperber (2000a) and Atran (2001), but I have shown that these are poor arguments and in the face of arguments for high fidelity copying we must reject the general Richerson and Boyd picture. Inaccurate averaging can suffice for evolutionary processes at the population level, but this is not what generally happens. Memes replicate.

5.4
Objection 4: There is no explanatory value of memes

Several critics have rejected memetics on the basis that it adds no new explanatory power to existing theories of cultural evolution (e.g. Edmonds 2005), that it is a trivial assertion that some ideas flourish while others die out, or that it is surplus to theories that explain culture fully by studying evolved psychology. As Richerson and Boyd have pointed out (2005, pg 250), the most important aspect of a theory is its productivity. What novel phenomena does memetics reliably predict? I will conclude this chapter's defense of memetics by addressing this issue and proposing a memetic index, which will quantify how likely a meme is to succeed. Such a measure of meme fitness will allow past explanation and future prediction of cultural phenomena. We need to be able to say this practice or that piece of information will spread. Not just because it is true or beautiful (see Dennett 1995).

It could be argued that all that is explained by the fitness of memes are facts about human psychology, that memetics is merely another way of talking

about our rich cognitive mechanisms, but it will become apparent that there are features of memes that enhance their fitness which are environmental or ecological factors and hence not aspects of human psychology. Memetics is not just another way of looking at human psychology. The selective forces on memes include psychology, ecology, and competition with other cultural objects. Memetics is a way of unifying these features of cultural change and influence.

Meme Fitness: The Memetic Index

Memes must be fit in their environment in order to persist. This means that they must tend to be taken up by human psychological devices and must also tend to cause their own reproduction in order to persist beyond their extinction in a particular host. The following features of a meme make it more fit in this ecology of human psychology.

1. The meme is visible. In order to be copied the meme must have an outward signature. The more robust this outward signature is the more chance a meme has of being copied. Robustness may be high in such cases where a meme causes people to construct an artefact from which the meme can be inferred. Indeed, any meme that exists as an artefact may enhance its own reproduction because it is seen often. Similarly, the more attention grabbing or holding a meme is in a particular context then the more likelihood it has of spreading.
2. The artefact (if any) itself can be stable and durable, readily made from local physical materials, easy to reverse engineer, and have an error-tolerant design (Sterelny 2006a).
3. The meme is easy to remember. If a meme builds on previously held knowledge, or if the meme has some emotional connection, or if the meme reinforces what is already believed, or is familiar, or if the meme is compressible, or if it's expression physically resembles something

129

already understood, then it may be easier to remember (or, of course, confuse with what is already held). For example, a star is more symmetrical and therefore compressible than a squiggle. A barbed fish hook builds upon a plain one.

4. A meme has to be copyable. Therefore, memes that depend on rare physical capacities such as great height will fail to spread. Some memes are just too difficult for many people to copy successfully. Memes need to have a tendency to actually get copied. If the meme is skill-based then being able to be broken into discrete subroutines will help (Sterelny 2006a).

5. A successful meme will have relevance for a particular population. For example the practice of greasing the thread of running shoe spikes with Vaseline before screwing them into the soles will be important for sprinters and other track runners, but many people will fail to pay attention to this otherwise irrelevant behaviour.

6. Fitness will also include generation time. A meme that is seldom exhibited may not spread easily.

7. Context is going to be very important. A meme for pumpkin soup, if presented in Spanish to an English speaker is not going to succeed.

8. Sometimes there will be physical environment constraints. Some memes may require special materials, or not be possible in freezing conditions.

These features of memes collectively form a memetic index. Fully worked out this could be a quantitative measure of a meme's likelihood to persist in a particular population in a particular context. Such a tool could be used to explain phenomena past, but importantly could be used to predict future phenomena. Such predictions would be a good test of the meme hypothesis. For example, empirical data could be collected detailing current opinions and practices among the general population of how to best lower cholesterol. Analyzing the cooking methods, medication regimes, beliefs and so forth and

applying the memetic index to patterns of behaviour we could predict which behaviours will persist over time. A follow up study could be conducted at a later date and comparisons made. The memetic index of various cholesterol lowering practices and beliefs would depend on the source of information, the ease of actually copying the behaviour, the prevalence of advertising, the appeal of diets, and so on. It ought to be the case that the memes with the highest memetic index, not the most successful cholesterol lowering strategy actually persist. It is this predictive element that makes memetics valuable. Other theories of cultural evolution do not allow for us to carve a unit of culture free of the collective whole and ask 'will this cultural object spread or not?' Psychology and non-content biases may explain how information flows through a population, but something like the memetic index is needed to make future predictions about a specific cultural practice. As well as being explanatory and predictive, memetics offers us one more insight. This is the realization that 'cultural evolution is an autonomous process over which we exercise a limited amount of control' (Distin 2005).

Memetics does need to be buttressed with at least one other theory, however, and this is the theory of what it is about the human mind that accounts for the process which memes parasitise. I will argue in the next chapter that memes are representations of real patterns, but in order for us to recognize patterns we must have some sort of pattern recognition device that precedes our acquisition of memes. We also take an attitude toward each pattern. To use Sterelny's example, the meme for mountain bike riding exists among skateboarders but they don't practice the behaviour (2006a). So each pattern representation has some attitudinal state associated with it. These attitudes are not memes. If memes must be parasitic on thinking, then they cannot fundamentally explain minds (contra Dennett 1995). They result from pattern recognition and intentionality, which likely evolved for other purposes.

Chapter Six: How Many Memes are there?

Overview

I have just argued that meme theory makes sense in the face of several important objections. I have also argued that both the Sperber and Richerson and Boyd approaches to culture are each flawed in important ways. Sperber is wrong to deny that replication occurs and Richerson and Boyd are wrong to accept arguments for low fidelity cultural transmission. I now turn to explaining to what degree memetics accounts for human culture. I will argue that one of the fundamental functions of human cognition is pattern representation. This is because we exist in a social, psychological and ecological world in which important patterns exist. Recognition of these patterns and our actions that result determine our fitness. I will argue that memes are representations of real cultural patterns in the world. By first outlining a theory of the nature of patterns introduced by Dan Dennett (1991), I will argue that humans are evolved to take a *memetic stance*, which ensures that we mentally represent patterns in culture. These mental representations are memes and are replicated from mind to mind with high fidelity. This process pervades human culture.

6.1
The nature of patterns

Kate Distin (2005) comments on the possibility that memes may be patterns. Or at least that memes are representations of *patterns of behaviour*. She bases her approach on Byrne and Russon's (1998) hierarchical approach to imitation (see also Minsky (1974) for a hierarchical approach to representing knowledge). However, if memes are representations of patterns then we need not limit ourselves to them being representations of patterns of behaviour, any representation of a pattern that tends to ensure its own replication could be a

meme. This may include the patterns that underlie institutions, artefacts, or their functions.

In chapter four, I left open the question of memes' existence. I suggested that ideally for memetics there would be a simple 'common format' that memes took in order to preserve their information. I return now to these questions. In order for me to illustrate how memes might exist I must first explore the notion of a pattern, and in particular I will now spend some time discussing the nature of pattern.

Bar-Code

Figure 6.1
(from Dennett 1991)

Figure 6.1 shows five examples of a pattern (A – E). We can call five black squares in a row 'bar-code' as Dennett does (1991). Note that each of the six patterns above has been created as bar-code plus some noise, but bar code is only visible in the first five.

A = Bar-code +25% noise
B = +10%
C = +25%

D = +1%

E = +33%

F = +50% (note bar-code is indiscernible in F)

Note that in A-E, bar-code can be discerned and reconstituted at a glance. That is, even with 33% noise we recognize bar-code as an idealized pattern. Also note that C and A both have 25% noise, and are different bitmaps, but the same pattern exists.

I say *exists*. But, is bar-code really there in A? Of course it is; you can see it with your own eyes! I'm not being facetious. We humans can pick up on incomplete or noisy data and abstract a real pattern. Bar-code is a pattern if there is data generated in non-random fashion that can be reliably abstracted from the noise. Humans are particularly good at picking up on patterns from noise. In fact we seem predisposed to pick up on pseudo-patterns when all there is is noise. Think of the gambler at the roulette wheel, think of superstitions about lucky sporting items. We are constantly on the look out for patterns. This ability is sensitive but not specific. Often we will see patterns that don't exist, such as in static on a TV screen, but this means that we have the ability to see most of those that do.

So the first point I wish to make is that human minds can consistently and usefully abstract patterns from noise, and abstract objects (such as the five black squares above) from messy data.

The Game of Life

Next, Dennett describes John Conway's 'Game of Life'. This is a single rule system for determining whether squares on a grid are lit or unlit. The rule applicable to every cell in the grid is as follows:

IF exactly two of a cell's neighbours are lit then stay in present state. IF exactly three of the neighbours are lit then go on. ELSE turn off.

When operating, this system may be fully described by generating a table that lists each square's Cartesian reference at time = n and denotes it a '1' for on or a '0' for off. However, when the iterating system is viewed on-screen some interesting phenomena occur. Different stable and/or predictable patterns occur. Four 'on' grid squares in a square remain on indefinitely. Other particular patterns appear to migrate across the screen remaining intact (gliders), or appear to eat each other (eaters), or blow smoke (puffers). There are functional-level regularities that are invisible to the observer of the list of coordinates that fully describes the physical state of the system. Note that at the level of physical description the system exhibits no 'movement'.

At 'level-one' on the other hand we can watch gliders get eaten by eaters and puffers blow smoke through rocks, and so on. This higher level can be described in a different language. We have *persisting objects* at this higher level.

It is theoretically possible to instantiate a universal Turing machine on the Life-world. Indeed Rendell (2001) has done this. This machine could run programs that any computer could run, such as a chess program for instance. We could then interpret an even higher level of pattern and predict its behaviour *as a chess player*. Note that the terminology we would use changes again. This is no longer a visual pattern it is an 'intellectual pattern'. Note also, that at this level prediction of the pattern can become fuzzy for epistemic reasons. Two different agents could predict different futures 'he'll take the bishop', 'no he'll protect his queen', even though the life-world is in relevant senses fully deterministic. Compare this to level-one where it is easy to predict that gliders always move down and across without exception.

With Conway's Game of Life in mind I can now make my second point about patterns, which is that there can be a hierarchy of patterns within a system. Some patterns clearly existing at high levels of description are invisible beneath.

Earlier we saw bar-code, a pattern created according to a 'hard-edged' process, ten columns of black dots followed by ten columns of white dots, in some cases obscured by noise. What if the pattern is created by the normal distribution of pixels around evenly spaced means, followed by a 'selective' slit-window applied iteratively which focuses pixels within its bounds? (see figure 6.2). Note that this process *ends* in bar-code (see the bottom row, or 3rd iteration, in 'slit window') rather than beginning with it. The pattern emerges over time. Still, 'bar-code' is there, existing, in every frame.

Slit Window

Figure 6.2
(from Dennett 1991)

The point here is that if a pattern is acted upon by a selective force then outliers can be discarded and the pattern can be refined and change over time. Similarly, if we are thinking 'bar-code' (with x % noise) as we look at this

sequence of patterns then we can easily remove the noisy outliers and even fill in missing bits of our original 'bar-code + x % noise' pattern in order that it more closely approaches pure 'bar-code'. A pattern (bar-code + x % noise) can cause us to abstract a pure pattern (bar-code) and one pattern (bar-code) can be used as a refining slit-window and applied to another pattern.

Parallel Patterns

The two elephants in figure 6.3 are patterns in two quite different bit-maps. They represent two quite different data sets but the interpretation of them (at a particular grain of analysis) is exactly the same. The two patterns are *functionally* the same. We all easily agree upon what the pattern is.

But note also the boxed sub-pattern, 'trunk'. Patterns can be composed of many other patterns. They can also form parts of larger patterns. In this case the pattern that is there is 'elephant', but it is composed of the sub-patterns 'trunk', 'leg', and so on.

Figure 6.3
(from Dennett 1991)

137

Before moving on to explain my view that social transmission is pattern representation replication, I need to make two last points. First, patterns need not be visual. There are musical patterns, patterns of behaviour, patterns of space and form, patterns in time, patterns of proprioception and so on. The examples I have offered so far of real patterns have been simple and visual for illustrative purposes. But we can appreciate patterns in other modalities. If we hear a musical scale, even if the expression is noisy, perhaps a note is missed, we can mentally represent the intended scale, the abstract pattern as opposed to the actual expression. As another example, there may be a traffic jam at a particular set of lights on average once a day. But if it occurs twice on one particular day, this does not alter the pattern that it occurs once a day. Patterns can be perceived not only visually, but in all aspects of social and cultural worlds. The second point is that sometimes patterns can have two different notation systems and still be the same pattern. We could imagine Conway's life-world as described by separate coordinate grids with the same underlying patterns existing at level-one. Similarly, temperature can be 0 C or 32 F with the same underlying pattern of molecular kinetic energy really existing. Kate Distin (2005) argues that humans have several different representational systems for encoding culture. We have language, symbolic drawing, mathematical notation, and so on. I would argue that there are patterns in human culture and these patterns are real and can persist regardless of which notation or representational system is used to describe them.

Here is an example of a non-visual pattern. Pneumonia, as a disease itself, is a pattern. Particular patients may exhibit any number of an array of signs. These could include fever, cough, sputum production, difficulty breathing, a racing heart, thirst, chest pain and so on. At some point the pattern resolves and we attribute pneumonia to the patient. The 'bit-map' in each case of pneumonia is likely to be different, just like the two elephants above, but the same identifiable pattern exists. The particular causative organism can even be different in each case.

Overall I am claiming that we see patterns, that we can see which data points out there in the world will fit a particular pattern, that patterns exhibit hierarchy and compositionality, and that patterns can be composed of other patterns. In the case of memes we see behaviours, or artefacts, or social institutions, and we recognize their patterns. The data points might be particular motor sequences, or types of materials, or goals an object is used to attain, or qualities of an object, or arrangements of people, or linguistic utterances, or sequences of symbols. Collections of these data points constitute real patterns in the world that we are experts at identifying and representing mentally.

6.2
Memes as real pattern representations

I am arguing for a pattern-based representational system for memes. There are real patterns in the cultural world and mental representations of real patterns are memes. Pattern representations are expressed as behaviour, artefacts, or institutions and these are the memetic phenotypes. Patterns are abstract idealizations of information in the world. Cultural transmission relies on *forming a pattern representation in one mind that is a replicate of a pattern representation in another mind on the basis of information passed through a public representation*.

This is somewhat confusing. But there are two types of entity involved. First there is a mental representation in person A's head. This is ultimately replicated in person B's head. But replication is via the intermediate real pattern of a public representation. We do not (and cannot) replicate a mental representation that is in someone's head without there being a public expression of this representation, whether in action or material culture. There is no cultural equivalent of DNA unwinding to give the replication enzymes

direct access to the code. But this does not matter as we have already established that indirect information copying ought to be seen as replication.

This copying process is analogous to the DNA-protein-DNA replication discussed in the previous chapter. The hypothetical enzyme reverse translatase cannot replicate a strand of DNA without there being a protein expression of the information in that DNA. Patterns of meme expression fill this role. The person we are copying the pattern representation from need not even be present if the expression of their representation is. All that we need is a mechanism of copying the mental pattern representation from its expression. This is the cultural equivalent of *reverse translatase*. This mechanism may be called the *memetic stance* and I will discuss this after giving a few examples of cultural patterns.

All culture can be described as patterns. All these patterns can be identified and mentally represented. I will describe what the data points are for the patterns 'dinner party', and 'pumpkin soup'. These comprise a cultural practice and an artefact. Just like the elephants in the diagram above are composed of pixels that carry information allowing us to see the pattern, 'dinner party', as a pattern, is composed of pixels of information. These include: Many people in one room, music coming from a stereo, glasses of wine, soup, a large table, lots of conversation, invitations, and so on. Dinner parties (and other cultural practices) occur because someone holds a mental representation of this pattern. By observing the expression of this meme, someone else can recognize the information that comprises this pattern. They can then form a mental representation of the same pattern. If the meme is a fit one it will cause further instances of mental representations of the same pattern via expression of the pattern in the world. If the meme appeals to what I *value* (say, company and drunkenness) then I will express the pattern publicly.

In addition to this, no two dinner parties are exactly the same. Noise within reasonable limits can be eliminated at any step if the pattern is recognised. If the observer is a 'pattern recognizer' then noise is no hindrance. Noisy patterns can be copied as idealizations. Dennett's analysis of patterns shows us that it is not impossible, or even difficult to extract the underlying nature of a pattern despite the noise in any particular instance of it, or, in the case of memes, in any particular phenotypic expression of the meme.

Now take pumpkin soup. Just like bar-code is made of five black squares separated by four white squares, actual pumpkin soup is made of pumpkins, milk, salt, and so on. The meme 'pumpkin soup' is a mental representation that includes data such as the following: 'pumpkin soup is that thing that looks yellow, feels hot, exists in bowls, consists of salt, pumpkins, pepper, is produced by cutting up pumpkins, is eaten before dinner, tastes yummy, etc'. The meme, if it is a fit one, causes the expression of a pattern in the world that causes a copy of the pattern representation to occur in someone else's mind. Just like each instance of bar-code is noisy, so each particular soup may use subtly different ingredients, or be served in a different sort of bowl, or blended in a different blender. However, the existence of bowls, blenders and pumpkins in a particular population means that noise can be minimized and constructive processes in any given instance of creating the real world pattern pumpkin soup will converge. Replication will be high fidelity.

Pumpkin soup is also a sub-pattern of the pattern 'dinner party', just like the elephant's trunk is a sub-pattern of the pattern 'elephant'. Because of the compositional, hierarchical nature of patterns we can copy parts of patterns and interchange them. Tomato soup can be substituted for pumpkin soup. Opium pipes can be substituted for wine glasses. But there comes a point where the pattern becomes indeterminate and cannot be restored to the idealization 'dinner party'. Forming a mental representation of the *new* pattern is how mutation can occur in memes.

Susan Blackmore makes the distinction between copying the product and copying the instructions (2000). For example, if I watch you make soup then make it myself later, I have copied the product. If, however, I follow a written recipe I am copying instructions. I believe this is the wrong way of looking at things. The written recipe itself is *also a product*, so, too, would be verbal instructions. These are different notation systems for the same pattern. Dennett explains how the distribution of the pixels of 'bar-code + noise' could be represented as four-bit hexadecimal chunks. If presented with this notation we would not see the pattern because our pattern recognition devices don't read hexadecimal. However, the pattern still exists (see the definition of a pattern above). The true instructions are mental representations. These representations store the pattern. When I copy you making soup, or copy the written recipe, I am copying the pattern of information. The recipe (and its associated cookbook photographs, *coupled with* my background knowledge) tell me that the pattern is: 'pumpkin soup is that thing that looks yellow, feels hot, exists in bowls, consists of salt, pumpkins, pepper, is produced by cutting up pumpkins, is eaten before dinner, tastes yummy, etc' There is a fundamental asymmetry here. Neither the actual soup nor its written recipe can exist without the mental representation 'pumpkin soup'. But this cultural object can exist without there being any instances of recipes or soups in the world. It is this representation that both soups and recipes have in common.

It must be noted at this point that natural kinds can also be represented as patterns. For example, a shrub is a pattern of sense data including leaves, green, short stature, bushy texture, and so on. We can see instances of shrubs as examples of this pattern. There are determinate and indeterminate instances of shrubs. We probably represent natural kinds in a similar way to cultural objects. However something is cultural only if there is transmission of the pattern to other minds. In that sense natural kind representations can be memes

if there is behaviour that clearly expresses publicly the information contained in the agent's mental representation of a natural kind.

As long as minds sustain them, pattern representations persist across time and could be acted upon by natural selection. The environment that this selection occurs in is largely an environment of human psychology. Patterns may be fit in themselves (e.g. easy to discern from their expression, or have attention grabbing content) or fit for the agent inferring them (good ideas).

6.3
The memetic stance

I have suggested that we are good pattern recognizers and that our skill in this domain allows us to extract the complex pattern, or meme, that someone must be mentally representing, from their expression of that meme. I now explain how this skill works and why we possess it. This skill may be called the *memetic stance* and it is the cultural analogue of the imaginary enzyme reverse transcriptase. It will become clear that there are two processes occurring in cultural transmission. Sometimes we see cultural patterns in the world on our own, and sometimes we are shown cultural patterns in the world. Hybrid cases obviously occur. I will now describe how a variety of mechanisms collaborate to ensure that we see patterns, and that there are also several mechanisms that ensure we can be shown patterns. These allow us to then represent the patterns and express them as behaviour, or, indeed, to teach them.

Dennett (1987) discusses the three 'stances' which one can take when interpreting a system. The *physical stance* is an interpretation based on deterministic laws of cause and effect. We can interpret and predict any system based on its physical properties. I know that a ball will roll downhill because of what I know about the physics of medium sized objects on Earth. We can also take the *design stance* and predict outcomes based on design. My alarm clock

(regardless of the physics involved) will beep at 8am because that is what it is designed to do. Finally, we can take the *intentional stance* and attribute beliefs and desires to a system (or agent) that we think it ought to have. These intentional states allow us to predict what the agent will do.

Dennett discusses our pattern recognition abilities with respect to intentional states (1991). He argues that we are exceptionally good at predicting and tracking the beliefs and desires of other agents. This is the cornerstone of folk psychology. We have evolved the ability to interpret the minds of other agents, not just humans, but animals, and even in some sense physical objects (Dennett 1987, Sperber 2000b, Bogdan 1997). This ability is clear when we realize that every time we cross the road at a red light, we do so knowing that the driver of an oncoming car intends to slow down and stop. There is no way that a coherent society as complex as ours could develop without the ability to reliably interpret other minds. This ability may be called folk psychology, 'mind-reading' in cognitive psychology, or to take Dennett's philosophical term, the *intentional stance*. We have evolved this ability because it is a fitness advantage to be able to predict what other agents are going to do. This confers a competitive, but also co-operative advantage. And our ability to interpret minds is, to put it mildly, bloody impressive.

Here's a famous example of just how impressive this ability is. Imagine a prediction contest between a humble Earthling and what Dennett calls a Laplacean Martian (1987). This is a creature so in tune with the physical world that it can calculate the position, momentum, and future effects of all atomic and sub-atomic particles it chooses to, and can compute such predictions far into the future. The Martian and the Earthling both witness the following scene. A woman picks up the telephone and says, 'Hello honey. Tonight? The boss? Well you better pick up the wine on your way home.' The Martian proceeds to calculate the position of each carbon atom in the plastic of the object the woman is holding to her ear. He works out the position of every

electron in the vicinity, the magnitude of the sound waves coming out of the object, and so forth. He predicts that a large metallic object will unload two organic creatures outside in about fifteen minutes and that one of the creatures will be in possession of a glass cylinder full of an alcoholic compound. The Earthling predicts that the woman's husband and his boss will arrive with some wine for a dinner party. The Martian is astounded. Given the Earthling's naïve epistemic relation to the physical particles of the world how could he have predicted the exact same thing as the Martian?

The answer of course is that the Martian was stuck using what Dennett calls the physical stance, and while his poor ectoplasm or silicon neurons or whatever were at the point of melting due to the computational complexity of the task, the Earthling took the intentional stance. The Earthling can do this because he is designed by natural selection to pick up on the patterns in the minds of other minded agents.

I propose another stance to add to Dennett's first three. Along with the physical, design and intentional stances we can add the *memetic stance*. The hardware of the memetic stance permits us to form representations of the cultural patterns of behaviour and artefacts all around us. This is just like the way reverse translatase would form a representation of DNA(1), in the form of DNA(2) by interacting with the protein that DNA(1) codes for. Just as the intentional stance can deduce patterns of intentionality from the impoverished signals we observe, so the memetic stance can form the pattern representation that underlies behaviour and tool making that we experience. Humans are expert meme manipulators, we can see patterns in the cultural world around us. The skills we possess in recognizing patterns cause the indirect replication of mental representations.

These skills are comprised of several different psychological devices, and are influenced by material culture and the behaviour of other individuals. There

are two broad processes occurring. At one end of a spectrum we have recognition of cultural patterns by the individual. At the other end we have direct teaching of patterns. In between various mechanisms facilitate the catalysis of pattern recognition. We possess a pattern recognition device, this has been clearly demonstrated by the examples of pattern recognition I have given above. We also have a constructed niche and ecological constraints which provide various background constraints upon what patterns we recognize. These constraints help to overcome the ambiguity of multiple patterns being equally applicable to observed phenomena. To first use a non-cultural example, if someone comes inside and they are wet, then it may either be the case that it is raining and they had no umbrella, or that there is a carwash blocking the footpath and the individual dashed through the hose. The signal 'person is wet' could represent either the situation 'is raining', or the situation 'car wash'. In cultural transmission the signal 'wearing a suit on Sunday' could represent 'going to work' or 'going to church'. However, in the case of cultural transmission there are important shared background constraints, which lead us to usually realize the correct set of information. In the case of 'wearing a suit on Sunday' we form the pattern 'going to church' because we don't tend to work on Sundays. As I have stated before, it is still possible that the pattern we form is the wrong one. We may see the bit-map 'elephant' but copy it as 'tiger'. In this case the memetic lineage ends and a new one begins.

The memetic stance also depends upon our theory of mind abilities. We can attribute beliefs and desires to agents on the basis of their behaviour and these beliefs and desires become data points in cultural patterns. This combination of pattern recognition, background constraints and theory of mind allow individuals to recognize patterns unaided. But, sometimes there is catalysis of the pattern recognition process. Culture that exists can permit agents to take a best guess at the pattern, then attempt to express the pattern as phenotype, and

then compare their resulting phenotype with the extant version. This is best illustrated with material culture.

Material Culture

In biology, DNA contains the gene for the protein DNA polymerase. The protein is the phenotypic expression of the information in the gene. This protein then plays a crucial role in the replication machinery of DNA. In social worlds material culture, as an expression of a mental representation, can play a crucial error-correcting role in the replication of memes. For example, I may try to copy my fellow and build a fish-trap. I can infer the pattern, from a combination of seeing him build a fish-trap, from my knowledge of fish and hunting, from what I know about rope and wire. I form a mental representation and may then try to build a trap but fail. It doesn't work correctly, the fish keep escaping. But then by consulting again the original artefact I can make the appropriate changes and edit my meme. The artefact is not a meme, but it is an editing device that can increase fidelity of my copying of the cultural object, the abstract pattern that I mentally represent.

Niche construction in general can have this effect. As Andy Clark points out, 'we have created a world so smart' (2003, pg 5). In this example there is hybrid learning occurring to a degree. This is where I am learning part of the solution from my friend's artefact and behaviour, and part of the solution by my own trial and error. Such a process can certainly accelerate learning but is the transmission of the fish-trap pattern memetic? I argue it is. When I first notice my friend with his fish trap I form a coarse grained pattern of what a fish trap is and how I might go about building one. There are important parts of this pattern that are lacking or incomplete. How did he bend those wires into a funnel like that? Where do I get thick flax for tying the lid securely? But there are constraints on the ways I can fill in this pattern. Perhaps only a certain sort of flax grows in this area, or more usefully, perhaps only one type is *cultivated*

here. If we are familiar with pliers in my village then this is the tool I will likely turn to for bending wire precisely. Perhaps the first time I make a fish trap it comes out like his + 25% noise. When I see his in action again I note the differences between his and mine and edit my next version. Now I have copied his trap + 10% noise. Again the copying event may be distributed through time. The more shared representations my friend and I have, the more chance I have of accurately copying the mental pattern he possesses. This meme will be fit in a population of flax-growing, plier-using fishermen, perhaps not in other populations.

The final process that occurs when taking the memetic stance is catalysis by others. This may take the form of shared attention, body language, resonance systems (see chapter two), sentimental bonding, topical prediction, specific language employed, or even the direct teaching of a pattern. Shared attention focuses the copier upon the pattern to be copied. Body language and resonance are feedback mechanisms, which signal when important aspects of the pattern have been recognized. Language and direct teaching are important aspects of pattern recognition, so I turn now to discuss them.

Language

Language contributes to the operational success of the memetic stance. When someone utters a statement, we need to distinguish between sentence meaning and speaker's meaning.

> Discovering sentence meaning is just a means to an end. Our true interest is in speaker's meaning. A speaker's meaning is a mental representation entertained by the speaker which she intends the hearer to recognize and to which she intends him to take some specific attitude (e.g. to accept as true). Verbal understanding consists in forming a metarepresentation of a representation of the speaker (Sperber 2000b).

Now, continues Sperber, sentences (which are public representations) typically *underdetermine* speaker's meaning. This is true, of sentences, but not of the beliefs and desires that we attribute to a speaker in virtue of the sentences that they utter (plus information from a host of other sources). Beliefs and desires can be data points in cultural patterns. There is, indeed, a process of compression (into language) and extraction (into parts of attributed patterns) occurring in cultural transfer, but the information reinserted in extraction accurately reflects the original information thanks to the collateral factors. These collateral factors, many of which I've mentioned above, are why utterances alone don't tend to encode memes. If getting speaker's meaning right is of high fitness value then there will have been selection for getting it right from limited evidence. For parts of meme patterns transmitted through language we ought to see high fidelity transmission of a concept-pattern. This is because we have evolved to communicate successfully through language. This can catalyze recognition of salient aspects of a real pattern in the cultural world. Language in any one mind is an accumulation of patterns that correspond to words, and these pattern-concepts must be similar in content from mind to mind evidenced by the very fact that we do successfully communicate complex ideas through language. Encountering words means that we tend to form pattern representations that match those of the speaker. It has to be emphasized that we are not copying the meme from the target's mind directly. We are forming a representation of a pattern, which is a replicate of the pattern representation in the target's mind, by virtue of information extracted from the expression or phenotype of the meme. The use of metaphor may be an example of this sort of thing. In metaphor difficult concepts can be transmitted by deliberately failing to achieve a certain speech effect, thus suggesting that the truly intended effect is in the vicinity of the overtly failed one (Sperber 2000b). Mechanisms like this all contribute to the success of the intentional and memetic stances.

Employing the memetic stance is a multifactorial process. We observe people's behaviour and then attribute a pattern that describes the behaviour, or the creations, or their ideas. This process is aided by hybrid learning mechanisms including conscious and unconscious social interaction, language and direct teaching. This memetic stance means that the represented pattern is the same pattern that is represented in the target's mind.

Of course, Sperber's complaint now arises, that we aren't *copying* anything. Is this true? In Godfrey-Smith's DNA and protein case *all* the information is still there (has been copied) at each stage it is just in a different format (protein rather than gene). It is true that the second DNA sequence may not be identical to the first due to code redundancy, but there are still tight constraints of what can be reverse transcribed, and a further protein copied from this second DNA sequence will still be the same as the previous protein. The situation is paralleled in cultural transmission.

It seems that we treat each other as intentional and memetic systems, and we *can't not* use the intentional and memetic stances. We *can't not* correctly infer the program of everyday cultural behaviours. We are *empowered* and *constrained* in these respects by our evolutionary history and our cultural context. To use Bogdan's catch phrase (1997) 'Minds mind minds'. We infer patterns from extant behaviours and attribute the programs. We do this automatically and with exceptional fidelity due to our evolutionary design. The memetic stance is ultimately what Sperber and Claidiere (2006) are trying to articulate when they discuss the general learning and imitation skills, and information and procedures, already present in human minds that allow us to construct memes from underdetermined signals of mental representation content. Some of Sperber's so-called cognitive attractors are what enables the memetic stance. Some cognitive hardware makes memetic transmission more likely. Richerson and Boyd, of course, argue a similar point when they model

strong attractors as tending to make discrete-replicating-trait dynamics models better[11].

I have argued that we have evolved to see patterns in the world. In fact we are so adept at this task that we often see patterns even when the data is random. This ability is reinforced when we are infants. Many of our games involve picking the odd one out, filling in the blank, completing the pattern. This ability may be a hindrance to successful cultural copying just as it is a hindrance to the gambler at the roulette wheel who thinks he can spot the pattern there. But I propose that it is our ability to see real (as opposed to pseudo) patterns that is the necessary pre-requisite for memetic transmission. It could even be this ability that intelligent non-cultural species lack. In order to see these cultural patterns we employ a strategy that relies on our ability to sense patterns, the constraints of our niche, our theory of mind, existing material culture, and social interaction with others. This interaction includes language, shared attention and pattern demonstration. These factors constitute a continuum from solo pattern recognition, to structured teaching of cultural patterns.

6.4
Concluding points

A common view of memes is that they are mental representations of cultural traits. But each instance or expression of a cultural trait differs so we must talk of memes as being representations of salient features of cultural traits. We represent the pattern not the noise. I have argued here that culture exists as patterns, and that memes exist as representations of real patterns, and are relevant for studies of human cultural evolution particularly with respect to their predictive value (as per the memetic index). But to what degree are they

---

[11] Of course we must remember that there are compelling reasons to believe that cognitive hardware is culturally encoded, see chapter three.

explanatory of culture? Even if cultural objects are passed among humans and form lineages, it seems that these lineages can sometimes blur. The closer the cultural inheritance follows the parent to child, G1 to G2 pattern, and the less peripheral noise there is, then the more obvious the phylogeny is. The more there is horizontal communication, the more 'cultural parents' there are, the more blurred and un-reconstructable material culture phylogenies become (Temkin and Eldredge 2007). One assumes this is the case for non-material cultural objects too. This is effectively Sperber's complaint. This would seem to indicate that the deeper we delve into cultural evolutionary time, the better memetics is as a model. The more that human communication and information sharing increases, the better population-based models may be for capturing the essence of cultural evolution.

Even if there are good arguments to divert our attention away from memes in the analysis of culture (and I have yet to be convinced by any), then it seems that some sort of pattern copying mechanism is at work and pervades cultural transmission. I suggest that once intentional thought and the ability to recognize patterns were in place in our ancestors' cognitive repertoire then memes were enabled. The newfound transmission of cultural objects necessitated the evolution of biases to shape this information. Memes co-evolved with genes and the dynamics of a twin inheritance system is *much* more complex than that of a single genetic replicator system. We will see dual inheritance, feed back and feed forward loops, opposing selection pressures, and counterintuitive dynamics.

Culture is composed of memes. Selfish memes have driven the emergence of psychological biases, and these biases are necessarily culturally encoded.

Chapter Seven: A Synthetic Approach to Culture

7.1 Summary of the discussion

I have argued that the origins of culture in our lineage necessitated the evolution of biases in order to tractably navigate the new information environment. I examined what the nature of these biases might be, concluding that they cannot be genetically coded to any significant degree. This is because of the flexibility they need to possess in the face of fluid cultural environments and because of the developmental mechanisms of the brain. I also outlined three possible views on what the nature of the information these biases act upon might be. First there was the view that cultural information is constructed and held in individual minds but does not flow in any meaningful replicative fashion between minds. Then I outlined the view that culture is information distributed in a population and cultural evolution is the temporal change of this population-level information as a result of low fidelity individual copying events. Finally, I outlined meme theory, which asserts that culture is usefully seen as bits of information that replicate in transmission.

In chapter two I outlined in broad strokes the deep origins of human social learning once such memes existed. Social learning allowed agents to save the costs of individual learning, but the need to avoid bad information, and to acquire good information, most likely caused the evolution of psychological biases that shape the flow of memes. The evolution and dynamics of some of these biases have been modelled mathematically by Richerson and Boyd, but their models are importantly thin. In particular it is possible to describe scenarios in which a particular given bias is maladaptive or acquires poor information for the agent. The biases are likely to be context-dependent and context-specific rather than genetic universals. An array of varying psychologies may well be the most adaptive situation for particular human

populations. Furthermore, if group selection is important in cultural dynamics then models based only on individual-level processes will need modifying.

In chapter three I explored the nature of these biases themselves. The biases are clearly important in human cultural transmission, but their nature is unclear. Part of the reason for this is the limited understanding we currently have of the development of the human mind-brain. Competing, predominantly nativist on one hand, and predominantly constructive on the other, models exist. I argued that there is no necessity that biases be genetically encoded, and, indeed, that this is unlikely given the rapid changes in humans' circumstances in recent times. I argued that the neurobiological evidence, and the changing requirements of human agents, suggests that the amount of bias hardwired into the developing brain must be minimal. Additionally, it seems possible that the information needed to establish useful heuristic biases could be culturally or niche encoded. This kind of plastic system for dealing with cultural information would be far more adaptive than the massively modular mind of traditional Evolutionary Psychology.

In chapters four and five I introduced and defended memetics as a coherent approach to culture. I determined that memes need to form identifiable lineages, to replicate with high fidelity, and to possess fitness values. I argued that memes do indeed meet these criteria. I noted particularly that meme fitness can be described according to a memetic index that denotes their fitness value in an ecology largely of human psychology.

Finally, in chapter six, I argued that human minds are pattern recognizing systems and that cultural information can be identified and transmitted as representations of real patterns. These 'cultural objects' are acquired by employing a memetic stance, which allows us to replicate mental representations of patterns by extracting information from public expressions of those representations. This process is catalysed by social interaction with

other agents. Humans' ability to recognize patterns and take an attitude toward them has enabled memes to flow through human cultures.

7.2 The synthetic view

I have argued that the memes view is *prima facie* correct and robust in the face of criticisms. The reason that this ought to be the preferred model of cultural transmission is because once we understand that replication events can be distributed in time then we see that much cultural transmission is high fidelity replication of mental representations of patterns in the world. These mental representations, as evolutionary replicators (because they satisfy the Dennett / Lewontin criteria for evolution by natural selection) have fitness properties of their own. These properties can be conceptualised with reference to the memetic index, which accounts for human psychology, ecological factors and other competing memes that compose the environment of memetic adaptadness. The fitness of memes is a predictive tool and also explanatory of much that is maladaptive for human agents. However, not all cultural transmission need be memetic. The existence of many memes does not rule out the possibility that sometimes agents do indeed average from a number of sources or blend cultural traits. In these cases meme lineages end and new ones begin. The process can be modelled using the population dynamic tools of Richerson and Boyd. Much of the Richerson / Boyd machinery can easily be applied to memetic transmission too, because, as these investigators note, in many cases low fidelity transmission dynamics can approximate those of replicators. This occurs when there are strong cognitive attractors that tend to draw information toward certain types of content.

Sperber argues that cognition and cognitive attractors underpin the transformations that occur in all cultural transmission events. I agree that there *is* such channelling of representations, but not for the reasons he thinks. There is no stable genetically evolved psychology. Sperber's view is conceptually

incoherent because of the unstable nature of the target that our minds must track. But each individual does have a stable psychology that reflects the social niche that they have grown up in. Memetic replication is stabilised by a mixture of cultural-group-dependent cognition, constructed human niches, and prior memes acquired by the agents. So in a sense the Sperber view applies to culture, we do have rich psychology, but this just means that memes are transmitted with even higher fidelity. This is the basis of the memetic stance.

Memetics comes with caveats. It is possible that the recent explosion of horizontal transmission in our lineage has indeed blurred the phylogenies of what were once distinct memes. Perhaps this means that memes are a better model of the deep origins of culture, or a better model in small cultural groups. But more empirical research is needed to confirm the extent of cultural phylogeny cross-fertilization. Such research would be in the vein of Temkin and Eldredge (2007).

So to summarize this synthetic view of human culture: first there was the emergence of social learning. Then our copying fidelity increased and memes appeared. The explosion in the number of memes necessitated biases in order to navigate the information environment. The biases are necessarily culturally encoded. Some of the Richerson / Boyd tools are likely to be useful in modelling cultural processes; however, the fundamental unit of culture is the meme. Memes are representations of real patterns in the world, and are replicators. We humans are pattern recognizers.

This synthetic theory can explain how humans are so adaptively fitted to their environments and yet at the same time appear frequently foolish. Examples of human cultural cleverness *and* maladaptation abound. For example, using hunting, shelter and clothing technologies, Inuit natives can survive in the Arctic Circle, but by chasing career goals, rather than reproductive ones, contemporary Westerners risk failing to replace populations.

The question at the outset of this thesis was, 'why do people?' We can see now that the answer is multifactorial. There are proximate causes of human behaviour, these include genetic causes, behaviour evoked by the physical environment, individual learning by trial and error, cultural practices transmitted vertically from parents, obliquely from the previous generation, and horizontally from contemporaries. Other proximate contributers to behaviour are psychological heuristics and biases which shape the uptake of memes, also, there may be content analysis of some pieces of cultural information. Finally, the ultimate cause of human behaviour is natural selection operating on all these proximate mechanisms. Such selection can target memes, individuals or groups. There can even be opposing selection at different levels and between substrates. A visual depiction of this array of influences on behaviour can be seen in appendix figure 2.

As Richerson and Boyd note, 'a certain frequency of maladaptation inevitably results from design trade-offs confronting an advanced cultural creature' (2005, pg 244). In the light of this thesis such maladaptation can be explained in at least four ways.

First, the maladaptive behaviour may once have been adaptive but has fallen into a fitness trap scenario. This occurs when a cultural trait originally conferred a relative advantage to individuals possessing it, but once every member of the population is practising the behaviour it ceases to confer relative fitness. One example of such a trait may be male combat to impress potential mates (wrestling, duels, rugby). When only a few males do it they are demonstrating their ability and willingness to fight. This is a fitness advantage in many situations. But, when everyone is doing it, it becomes mostly a source of injury and everyone's fitness decreases en masse.

Second, maladaptation may be the result of a fit but negative meme. If the memetic index of a meme is high then it will take hold in a population regardless of its effects on the survival of genetic replicators. For example, the practice of burning fossil fuels to generate electricity seems a candidate for a high memetic index. It is visible, common, relevant, requires no special materials (compared to, say, nuclear power), and so forth. But it seems that ultimately it may be maladaptive as the resulting global warming causes crops to fail and living environments to flood.

Thirdly, psychological biases may have driven the acquisition of a poor behaviour. If agents practice a prestige bias, they may tend to acquire behaviours that do not confer them any advantage. In the 1990s Dennis Rodman dressed in woman's clothes, dyed his hair red and pierced his body. This behaviour was profitable for him due to his role in the entertainment industry (as an NBA basketball player and TV show host), but if imitated by others may result in attack on a Friday night downtown.

Finally, a novel niche may have made a once adaptive behaviour maladaptive. Thirty years ago it was good to aspire to possess a big new car. This was a signal of material wealth (often a fitness correlate). But in a crowded inner city niche, in an era of rising fuel and parking prices, this behaviour is becoming maladaptive unless it is performed by the genuinely rich. Transitions from what was adaptive to what is maladaptive can occur in the space of a single generation, even within a generation if culture changes quickly enough.

Overall it seems that cultural flow is of high fidelity. Conformity (or similar biases) are not needed to provide redundancy, but rather are needed to combat selfish memes. Memes with a high memetic index can lead to the retention of cultural innovation. And success / prestige biases (or similar) can explain how memes that are fit for agents, but have low memetic indices, can persist, for

example the boiling of drinking water, which is neither appealing nor obviously relevant to those who aren't aware of the germ basis of disease.

7.3 Complexity

This whole issue, however, is formidably complex. The problems of modelling cultural systems are similar to those for neuroscience. Instead of 10-100 billion neurons, cultural systems have 6 billion people. Rather than trillions of dendritic and axonal connections we have trillions of inter-agent interactions. But it is not good enough to model this system with binary variables or with simplified algorithms. What needs to be studied are detailed and complex models of all the variables involved in cultural transmission (see appendix figure 2). We have much of the data available now, but it is strewn across disciplines as diverse as the humanities, social sciences, and sciences. The study of culture needs a unified approach working within a single paradigm. The best candidate for this paradigm is the extended evolutionary paradigm. Just as no *one* human can visualise the genetic information established by the human genome project, and just like no *one* cognitive scientist knows about how the entire brain works, no one researcher in the field of culture, its transmission and evolution can have the whole detailed picture. But working with a sketch of this picture in mind we can model its parts and stick these part models together if we can be precise about the outputs of each model. Of course we don't want to just construct a one-to-one map of human culture. This tells us nothing new about the processes involved. We want to understand what the essential and inessential elements of the processes are at a meaningful grain of analysis.

But in order to achieve this, then those who study culture must all be speaking the same language and working against a backdrop of the same paradigm. The best candidate for a framework that can unify anthropology, history, psychology, economics, sociology, and all the other cultural sciences is the

extended evolutionary paradigm (appendix figure 1). Only with this always in mind can we move on to understand memes, niches, genes, societies, institutions, human behaviour and psychology. The answers will resolve bit by bit, but disparate academics need to be working within one system. As it stands each discipline is asking questions about human culture but there is no integration. Psychologists may ask why an infant does such and such, and focus on development, but development is guided by memes, genes, and human social organisation. Sociologists may ask why a particular institution exists, but institutions result from the interactions of memes, niches, and cultural group selection.

Tinbergen discusses his four 'whys' of animal (read human) behaviour. We can answer the question of why we humans do anything by studying mechanisms, development, function (from an evolutionary point of view), or phylogeny (Dunbar and Barrett 2007). Answering each question is important and with all the answers we form a complete explanation of behaviour. But the answer to any one fits within the extended evolutionary framework, and for this fit to be seamless the framework needs to be understood and borne in mind. Studies of consumer behaviour, for example, make sense in this way when couched in terms of psychological biases, meme fitness and cultural niches. Economics, though interesting on its own, forms part of a richly interwoven story of why humans do what we do. Similarly, the analysis of heroic modern medicine from a sociological and political point of view is interesting, but when seen as a group fitness trap (as more and more resources get poured in for diminishing returns in the face of universal demands) longer term implications in an evolutionary framework become clear.

The synthetic evolutionary rules of culture and the extended evolutionary paradigm allow us to analyse the same data we always have, but in an integrated way, usefully rather than in disconnected, foundationless, single-discipline languages and folk terms.

7.4 Future research

If culture is one aspect of an extended Darwinian model and the synthetic thesis I have outlined here is approximately correct, then what are the most likely avenues for future research?

Memetics

The most pressing need for memetics is some sort of empirical framework that can be used in order to test the many hypotheses it has generated. I suggest that two things need to be done. First, a more fully worked out memetic index needs to be established from psychological and cultural data. If we can identify features of memes that survived in the past then we can start to test the hypothesis by making novel predictions. What we also need to try to identify are the factors that make one meme a cultural parent of another. In what circumstances are there identifiable cultural parents and how often do we really reconstruct from many sources at once. I have already suggested at least one way that such empirical work could proceed.

There is one more area where empirical work in memetics could proceed and this is to extend the work done looking at the power of non-content biases. I have suggested that the content of a meme can nullify the effect of non-content biases in certain pay-off games. There are many more avenues in which the relative forces of meme fitness and context biases could be compared.

The Information Environment

Given the array of situational factors influencing biases and their worth, we ought to find differences in biases *between* groups (contexts). Factors that make this likely include, the probable culturally-transmitted aspects of

psychological biases (see chapter three, also Driscoll 2008), also the different sorts of biases that are adaptive according to mathematical models of different circumstances. I have mentioned group density as probably important in driving the evolution of certain biases for example. For many of the same reasons we ought to find similarities in biases *within* groups. But we need to look carefully for these similarities and differences. It is possible that data already collected and supporting, for example, a universal prestige bias may actually demonstrate conformity in some cases, and prestige in others, due to the shared features of these two biases (see chapter two). Now that we know we have to look very carefully more useful data can be collected.

It would also be interesting to know what the relative contributions of different biases are in different circumstances. This is especially true where competing biases are at the agent's disposal. More ethnographic and psychological data in these areas can only help. And the information gained will enable us to construct mathematical models that more accurately reflect reality.

The Psychological Biases

A major issue is coding. If genes can only be seen as causal IF-THEN cascades, then where does the information lie? Is there any coding information at all? And if genes can't be said to 'code' for biases that shape cultural transmission, then what can? Given the central tenets of this thesis, I would expect to find cross-cultural differences in cognitive architecture for transmitting and taking up cultural information, where this emergent cognition depends on the niche. There ought also to be evidence of non-stationary learning in the brain that permits the tractable acquisition of biases without the need for genetic pre-specification. I have suggested a learning model in which representation space is added sequentially as the system learns. There is no end of empirical research that could be done looking at such things as the micro-neuroanatomy of human frontal cortex, at in utero brain activity, at proactive

brains learning without input, which would all help to confirm or reject this position. Examples, too, of actual mechanisms of non-stationary learning in neural nets, their possibilities and limits, will all contribute to growing understanding in this area and hence of the nature of the psychological biases and the social construction of cognition.

We can hypothesis test some of the theories of neural constructivism and the coding of cognitive mechanisms by examining many of the myriad psychological and psychiatric diseases. We can ask either, what would be the outcome of a developmental process in which X was an atypical environmental condition? Or conversely, can our theory of cognitive emergence explain these rare cases? Schizophrenia could conceivably be a useful avenue of research. It is a very prevalent disease and some data suggest it is a disease of modernity. Is there environmental coding for schizophrenia? Could human technologies and developmental environments have produced it? These questions must, of course, wait for another day.

Finally, many of the arguments in this thesis depend on certain features of the theories of group selection and niche construction. I have not defended these theories here but their relevance is clear when one looks at their important roles the extended evolutionary paradigm (appendix figure 1).

The main objections to this new kind of synthetic approach to human culture certainly include the following: doubting the reality of memes, arguing a central role for certain genetically coded, evolved, psychological biases in cultural transmission, and the nativist-selectionist view of mind. I hope that I have shed sufficient doubt on my opponents' views that their prevailing orthodoxy cannot be seen as dogma, and, indeed, may be wrong in many instances.

Appendix

Figure 1   **The Extended Evolutionary Paradigm**
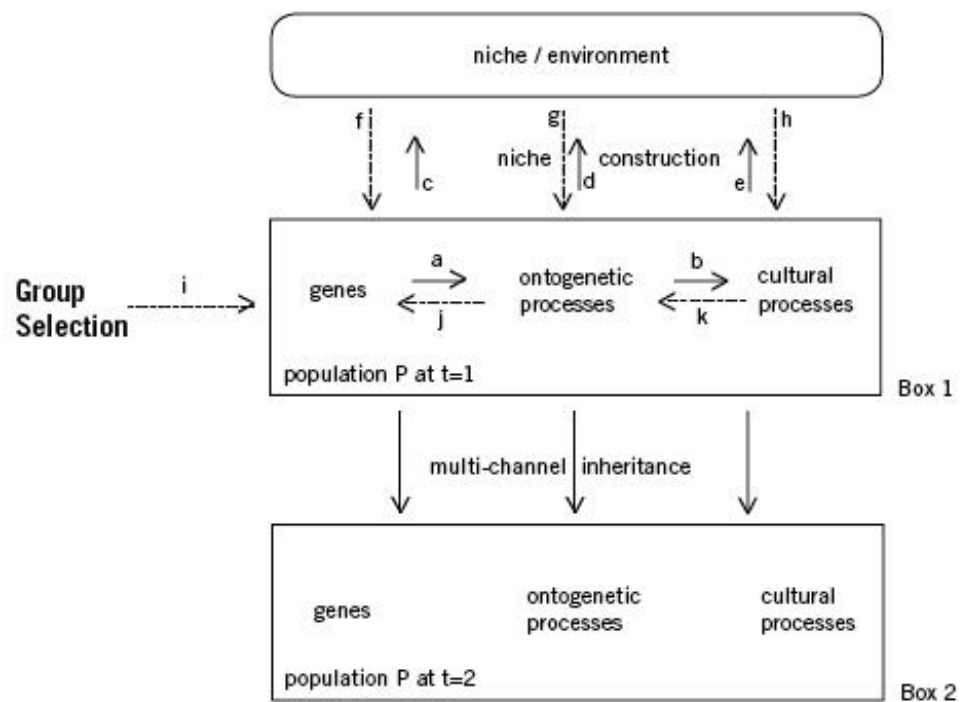
The extended evolutionary paradigm (after Laland et al. 2000). Note the traditional causal pathways (a,b) but also niche constructive processes (c,d,e) and feedback in the form of selection pressures on all three information systems (f,g,h,j,k). Note also, the role of group selection in moderating all these population-environment interactions (i). Box 1 represents the population at t = 1, box 2 demonstrates that multi-channel inheritance through all three information systems (genes, ontogenetic processes and culture) determines the composition of the population at t = 2.

Figure 2   Important Influences in Cultural Transmission and Human Behaviour
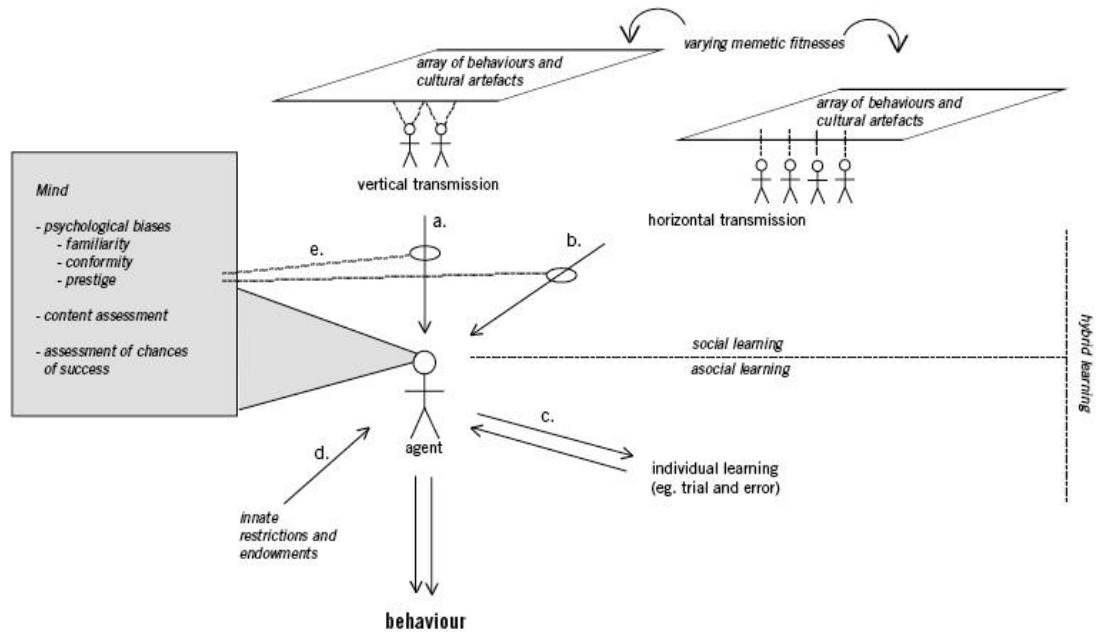
Important influences in cultural transmission and human behaviour. Note the multifactorial mediators of agent behaviour. The upper half of the figure illustrates social learning, composed of vertical (a, from G1) and horizontal (b, within G2) transmission process, filtered by cognitive mechanisms (e). The lower half of the diagram illustrates the influences on asocial agents. There can be individual learning by repeated trial and error, and there exist innate restrictions and endowments (which may of course affect what may be learned socially too). Taken together social learning processes and individual learning processes (guided by cognitive biases) enable the agent to learn by hybrid means.

References

Asch, S. (1955). Opinions and social pressure. *Scientific American,* 193: 31-35.

Atran, S. (2001). The trouble with memes: Inference versus imitation in cultural evolution. *Human Nature*, 12(4): 351-381.

Aunger, R. (ed) (2000). *Darwinizing Culture: The status of memetics as a science*. Oxford: Oxford University Press.

Aunger, R. (2007). Memes. In Dunbar and Barrett. *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Avital, E., and Jablonka, E. (2000). *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge: Cambridge University Press.

Bergstrom, B., Moehlmann, B., Boyer, P. (2006). Extending the Testimony Problem: Evaluating the Truth, Scope and Source of Cultural Information. *Child Development*, 77(3): 531-538.

Bjorklund, D., Hubertz, M., Reubens, A. (2004). Young children's arithmetic strategies in social context: How parents contribute to children's strategy development while playing games. *International Journal of Behavioural Development,* 28: 347-357.

Blackmore, S. (2000). The Memes' Eye View. In R Aunger (ed). *Darwinizing Culture: The Status of Memetics as a Science*. Oxford: Oxford University Press.

Blackmore, S. (2001). Evolution and Memes: The Human Brain as a Selective Imitation Device. *Cybernetics and Systems,* 32: 225-255.

Blackmore, S. (2005). A possible confusion between mimetic and memetic. In S Hurley, N Chater (eds). *Perspectives on Imitation: From mirror neurons to memes*. Cambridge MA: MIT Press.

Bloch, M. (2000). A Well Disposed Anthropologist's Problems with Memes. In R Aunger (ed). *Darwinizing Culture: The Status of Memetics as a Science*. Oxford: Oxford University Press.

Boase, J., Wellman, B. (2001). A Plague of Viruses: Biological, Computer and Marketing. *Current Sociology,* 49(6):39-55.

Bogdan, R. (1997). *Interpreting Minds.* Cambridge MA: MIT Press.

Bond, R., Smith, P. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgement task. *Psychological Bulletin*, 119: 111-137.

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

Byrne, R. (2005). Social Cognition: Imitation, Imitation, Imitation. *Current Biology,* 15(3):498-500.

Byrne, R., Russon, A. (1998). Learning by Imitation: A hierarchical approach. *Behavioural and Brain Sciences,* 21: 667-721.

Carr, M., Kurtz, B., Schneider, W., Turner, L., and Borkowski, J. (1989). Strategy acquisition and transfer among American and German children: Environmental influences on metacognitive development. *Developmental Psychology*, 25: 765-771.

Chaminade, T. (forthcoming). From mirror neurons to human cognition: the missing links.

Clark, A. (2001). Reason, Robots and the Extended Mind. *Mind and Language,* 16(2): 121-145.

Clark, A. (2003). *Natural Born Cyborgs*. Oxford: Oxford University Press.

Daley, T., Whaley, S., Sigman, M., Espinosa, M., & Neumann, C. (2003). IQ on the rise, the Flynn effect in rural Kenyan children. *Psychological Science,* 14(3): 215-219.

Dawkins, R. (1976, revised 2nd ed. 1989). *The Selfish Gene*. Oxford: Oxford University Press.

Dehaene-Lambertz and Dehaene. (1997). In defense of learning by selection: neurobiological and behavioural evidence revisited. *Behavioural and Brain Sciences*, 20: 560-561.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, Mass.: MIT Press.

Dennett, D. (1991). Real Patterns. *Journal of Philosophy,* LXXXVIII (1): 27-51.

Dennett, D. (1995). *Darwin's Dangerous Idea*. London: Penguin.

Dennett, D. (1998). *Brainchildren: Essays on designing minds*. Cambridge Mass.: MIT Press.

Dennett, D. (2000). Making tools for thinking. In D Sperber (ed). *Metarepresentation*. New York: Oxford University Press.

Diamond, J. (1997). Guns Germs and Steel: the fates of human societies. New York: W.W. Norton.

Distin, K. (2005). *The Selfish Meme*. Cambridge: Cambridge University Press.

Driscoll, C. (2008). The problem of adaptive individual choice in cultural evolution. *Biology and Philosophy*, 23(1): 101-113.

Dunbar, R., Barrett, L. (eds). (2007). *The Oxford handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Durham, W. H. (1991). *Coevolution: Genes, culture and human diversity*. Standford CA: Stanford University Press.

Eckel, C., Wilson, R. (2007). Social Learning in Coordination Games: Does Status Matter? *Experimental Economics* 10(3): 317-329.

Edmonds, B. (2005). The revealed poverty of the gene-meme analogy – why memetics per se has failed to produce substantive results. *Journal of Memetics – Evolutionary Models of Information Transmission*, 9. http://jom-emit.cfpm.org/2005/vol9/edmonds_b.html

Eigen, M., Schuster, P. (1977). The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften,* 64**:**541-565.

Everett, D. (2005) Cultural constraints on grammar and cognition in Piraha. *Current Anthropology*, 46:621-646.

Flynn, J. (1987). Massive IQ gains in 14 Nations: What IQ tests really measure. *Psychological Bulletin*, 101: 171-191.

Flynn, J. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54: 5-20.

Gigerenzer, G. and Todd, P. (1999) *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Gigerenzer, G. and Selton, R. (2001). *Bounded Rationality: the adaptive toolbox*. Cambridge, MA: MIT Press.

Gil-White, F. (2005). Common misunderstandings of memes (and genes): The promise and the limits of the genetic analogy to cultural transmission processes. In S. Hurley & N. Chater (eds). *Perspectives on imitation: From mirror neurons to memes*. Cambridge Mass.: MIT Press.

Godfrey-Smith, P. (2000). The replicator in retrospect. *Biology and Philosophy,* 15: 403-423.

Gold, E. (1967). Language identification in the limit. *Information and Control*, 10: 447-74.

Harris, P. (2002). Checking Our Sources, the origins of trust in testimony. *Studies in History and Philosophy of Science*, 33: 315-333.

Henrich, J., Boyd, R. (1998). The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behaviour*, 19: 215-241.

Henrich, J., Boyd, R. (2002). On modelling cognition and culture: why cultural evolution does not require the replication of representations. *Journal of Cognition and Culture*, 2(2): 87-112.

Henrich, J., Boyd, R. and Richerson, P. J. (forthcoming). Five Misunderstandings about Cultural Evolution, in D. Sperber (ed.) *Epidemiology of Ideas*. New York: Open Court Publishing.

Henrich, J., Gil-White, F. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behaviour*, 22: 165-196.

Henrich, J., McElreath, R. (2007). Dual Inheritance Theory: the evolution of human cultural capacities and cultural evolution. In Dunbar and Barrett (eds). *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Henshilwood, C., Marean, C. (2003). The Origin of Modern Human Behaviour: Critique of the models and their test implications. *Current Anthropology,* 44(5): 627-637.

Hewlett, B., Cavalli-Sforza, L. (1986). Cultural transmission among Aka Pygmies. *American Anthropologist*, 88(4): 922-934.

Jablonka, E., Lamb, M. (2005). *Evolution in Four Dimensions*. Cambridge, Mass.: MIT Press.

Keysers, C., Perrett, D. (2004). Demystifying social cognition: a Hebbian perspective. *TRENDS in Cognitive Science*, 8(11): 501-507.

Klintsova, A., Greenough, W. (1999). Synaptic Plasticity in Cortical Systems. *Curr Opin Neurobiol*, 9(2): 203-208.

Koenig, M., Fabrice, C., Harris, C. (2004). Trust in Testimony: Children's use of true and false statements. *Psychological Science*, 15: 694-698.

Kostovic, I., Jovanov-Milosevic, N. (2006). The development of cerebral connections during the first 20-45 weeks gestation. *Seminars in Fetal and Neonatal Medicine*, 11(6): 415-422.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lake, M. (1998). Digging for Memes: the role of material objects in cultural evolution. In: C. Renfrew & C. Scarre (eds). *Cognition and Material Culture: the Archaeology of Symbolic Storage*. McDonald Institute for Archaeological Research. Oakville: The David Brown Book Company.

Laland, K. (2007). Niche construction, human behavioural ecology and evolutionary psychology. In Dunbar and Barrett. *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Laland, K. (forthcoming). Animal Culture: Problems and solutions. In K. Laland & B. Galef (eds). *The Question of Animal Culture*. Harvard University Press.

Laland, K., Odling-Smee, J. (2000). The Evolution of the Meme. In R. Aunger. (ed) *Darwinizing Culture: the status of memetics as a science*. Oxford: Oxford University Press.

Laland, K., Odling-Smee, J., Feldman, M. (2000). Niche constrction, biological evolution, and cultural change. *Behavioural and Brain Sciences,* 23: 131-175.

Lewontin, R. (1980). Adaptation. Reprinted in R. Levins & R. Lewontin, *The Dialectical Biologist*. Cambridge: Havard University Press.

Lillard, A. (1998). Ethnopsychologies: cultural variations in theories of mind. *Psychological Bulletin*, 123: 3-32.

List, C. (2005) Group knowledge and group rationality: a judgment aggregation perspective. *Episteme,* 2(1): 25-38.

McElreath, R., Henrich, J. (2007). Modelling Cultural Evolution. In Dunbar and Barrett. *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Mameli, M. (2006). Understanding culture: a commentary on Richerson and Boyd's *Not By Genes Alone*. *Biology and Philosophy*. online first, 10 Nov 2006.

Mameli, M. (2007). Evolution and Psychology in Philosophical Perspective. In Dunbar and Barrett. *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Marcus, G. (2004). *The Birth of the Mind*. New York: Basic Books.

Maynard-Smith, J., Szathmary, E. (1995). *The Major Transitions in Evolution*. New York: W.H. Freeman.

Meltzoff, A., Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312): 74-78.

Mesoudi, A., Whiten, A., Laland, K. (2006). Towards a unified science of cultural evolution. *Behavioural and Brain Sciences,* 29: 329-383.

Miall, R. (2003). Connecting mirror neurons and forward models. *NeuroReport,* 14(16): 1-3.

Minsky, M. (1974). A Framework for Representing Knowledge. In, J. Haugeland (ed). *Mind Design II*. Cambridge Mass.: MIT Press.

Mithen, S. (1996). *The Prehistory of the Mind*. London: Thames and Hudson.

Montague, R. (2006). *Why Choose This Book: How we make decsions*. New York: Dutton.

Moss, L. (2003). *What Genes Can't Do*. Cambridge Mass.: MIT Press.

Mountcastle, V. (1978). An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System. In Edelman and Mountcastle (eds). *The Mindful Brain*. Cambridge Mass.: MIT Press.

Odling-Smee, J., Laland, K., Feldman, M. (2003). Niche *Construction: the neglected process in evolution*. Princeton NJ: Princeton University Press.

Pica, P., Lemer, C., Izard, V., Dehaene, S. (2004). Exact and Approximate arithmetic in an Amazonian Indigene group. *Science,* 306: 499-503.

Plotkin, H.C. (1996) Non-genetic transmission of information: Candidate cognitive processes and the evolution of culture. *Behavioral Processes,* 35: 207-213.

Prinz, J. (forthcoming). The return of concept empiricism. In H Cohen & C Leferbvre (eds). *Categorization and Cognitive Science*. Elsevier.

Quartz, S., Sejnowski, T. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioural and Brain Sciences*, 20: 537-596.

Ramachandran, V. (2000). Mirror neurons and imitation learning as the driving force behind 'the great leap forward' in human evolution. The Edge website (6/29/00): http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html

Rendell, P (2001). This is a Turing machine implemented in Conway's Game of Life. http://rendell-attic.org/gol/tm.htm. (Accessed 11/12/00).

Richerson, P., Boyd, R. (2005). *Not by Genes Alone: How culture transformed human evolution*. Chicago: University of Chicago Press.

Ridley, Mark. (2000). *Mendel's Demon: Gene justice and the complexity of life*. London: Weidenfeld and Nicolson.

Ridley, Matt. (2003). *Nature via Nurture*. London: Fourth Estate.

Rogers, E. (1995). *Diffusion of Innovations*. New York: Free Press.

Samuels, R. (1998). What Brains Won't Tell Us About the Mind: A critique of the neurobiological argument against representational nativism. *Mind and Language*, 13(4): 548-570.

Sarnecki, J. (2006). Developmental objections to evolutionary modularity. *Biology and Philosophy*, 22(4): 529-546.

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). *The influence of culture on visual perception*. Indianapolis: The Bobbs-Merrill Company, Inc.

Shanahan, M. (2004). The Frame Problem. *Stanford Encyclopaedia of Philosophy*. http://plato.stanford.edu/entries/frame-problem/ (accessed 13/1107).

Shennan, S. (2006). *Not By Genes Alone: How Culture Transformed Human Evolution*, by Peter J. Richerson and Robert Boyd. *Biology and Philosophy*, online first 18 Oct 2006.

Sloan-Wilson, D. (2007). Group-level Evolutionary Processes. In Dunbar and Barrett. *The Oxford Handbook of Evolutionary Psychology*. Oxford: Oxford University Press.

Sperber, D. (1996). *Explaining Culture*. Oxford: Blackwell Publishers Ltd.

Sperber, D. (2000a). An Objection to the Memetic Approach to Culture. In R Aunger (ed). *Darwinizing Culture: the status of memetics as a science*. Oxford: Oxford University Press.

Sperber, D. (2000b). Metarepresentation in an Evolutionary Perspective. In D Sperber (ed). *Metarepresentation: A Multidisciplinary Perspective*. Oxford: Oxford University Press.

Sperber, D., Claidiere, N. (2006). Defining and Explaining Culture (comments on Richerson and Boyd, *Not by Genes Alone). Biology and Philosophy*, online first 25 May 2006.

Sterelny, K. (2003a). *Thought in a Hostile World*. Oxford: Blackwell.

Sterelny, K. (2006a). Memes Revisited. *Brit. J. Phil. Sci*, 57: 145-165.

Sterelny, K. (2006b). The Evolution and Evolvability of Culture. *Mind and Language*, 21(2): 137-165.

Sterelny, K. (2007). SNAFUs: An Evolutionary Perspective. *Biological Theory* 2(3): 317-328.

Suroweicki, J. (2004). *The Wisdom of Crowds*. London: Abacus.

Taillard, M. (2004). Adaptive persuasion. *UCL Working Papers in Linguistics,* 16: 247-263.

Taylor, T. (2006). The human brain is a cultural artefact. In J Brockman (ed). *What is your dangerous idea?* London: Simon and Schuster.

Temkin, I., Eldredge, N. (2007). Phylogenetics and Material Cultural Evolution. *Current Anthropology,* 48(1): 146-153.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, Mass.: Harvard University Press.

Tooby, J., Cosmides, L. (2000). Consider the Source: the evolution of adaptations for decoupling and metarepresentation. In D Sperber (ed). *Metarepresentation*. New York: Oxford University Press.

Vinden, P. (1999). Children's understanding of mind and emotion: A multi-culture study. *Cognition and Emotion*, 13(1): 19-48.