The Evolution and Evolvability of Culture

October, 2003

Kim Sterelny
Philosophy Program, Research School of the Social Sciences, Australian National University, and,
Philosophy Program, Victoria University of Wellington
e-mail: Kim.Sterelny@vuw.ac.nz or kimbo@coombs.anu.edu.au

To appear in Dennis Walsh (ed) Twenty Five Years of Spandrels, Oxford University Press

I The Accumulation of Cognitive Capital

Joseph Henrich and Richard McElreath begin their survey of theories of cultural evolution with a striking historical example. They contrast the fate of the Bourke and Wills expedition — an attempt to explore some of the arid areas of inland Australia — with the routine survival of the local aboriginals in exactly the same area. That expedition ended in failure and death, despite the fact that it was well equipped, and despite the fact that those on the expedition were tough and experienced. For survival in such areas depended on accumulated local knowledge. The locals had learned how detoxify seeds before making bread from them, and how to catch the local fish. Bourke and Wills and their companions lacked this local knowledge, and died as a result (Henrich and McElreath 2003).

This example is far from exceptional. In their writings on culture and evolution, Peter Richerson and Robert Boyd invite their readers to consider their bleak prospects if given the task of surviving without modern technology on the Altar Desert's "The Devil's Road" between Mexico and Old California. A lot of tough and experienced folk did not make it across that road. But the Papago Indians managed to make a living there on the basis of stone, wood and bone equipment and a lot of local knowledge (Boyd and Richerson 2000). Similar examples can be multiplied indefinitely, and they make a most important point. Humans are often very well adapted to their local environment. The behavioural adaptations of, say, the seal and whale hunting Inuit, in making kayaks and weapons out of the very restricted range of materials available to them in the arctic north, are as impressive as the morphological adaptations of the creatures they hunted. Moreover, that adaptation is knowledge-intensive: appropriate action in the artic north, in the altar desert, or in central Australia depends on the accumulation of extensive cognitive capital.

The existence and importance of cognitive capital, in turn, makes evolutionary approaches to culture plausible. For cognitive capital is not accumulated rapidly. It is surely likely that the information of aboriginal and indian foragers was accumulated piecemeal and over time: it was the work of generations rather than genius. Moreover, it is very likely that the accumulation mechanism is sensitive to actual success rather than to the anticipation of success[1]. We can reasonably conjecture that the informational bases of Australian aboriginal desert survival was built by the retention and faithful transmission of successful experiments; by recognising and responding to what worked rather than (exclusively) by recognising in advance what would work.

Thus in some form selection-like processes have played an important role in building human cultural adaptations to their local circumstances. One way to interpret Gould and Lewontin's Spandrel's paper (and much of their individual work) is to see it as a set of theses about the evolvability of the biological world. As they saw it, the neoDarwinian synthesis took the potential for evolutionary change to be very rich. In particular, it was not constrained by limits on variation. The supply of variation to selection was dense and isotropic. That is to say, the supply mechanisms tend to produce variants in each trait of the organism on both sides of its current mean, densely, and independently of variation in other traits (Gould and Lewontin 1978; Lewontin 1985; Gould 2002). Neither Gould nor Lewontin denied the importance of adaptation, but they regarded these views on evolvability as extremely suspect. On their view, the supply of variation to selection is not typically dense and isotropic. Many of the fundamental features of the tree of life are explained by constraints on variation, not by selection. The fate of Gould and Lewontin's much more constrained conception of the evolvability of the biological world is still very much open (Sterelny 2000; Sterelny 2003a). My aim in this paper is to pursue an analogous set of questions about evolvability and human cultures. Under what conditions are selection-like processes powerful, and how widely have those conditions characterised human cultural environments? To what extent is that flow constrained, and by what? These issues have been canvassed in an increasingly impressive way by Peter Richerson, Robert Boyd and their co-workers[2] and I shall exploit their corpus in exploring these issues.

---

[1] Though variation very likely is not wholly blind: innovations in existing technique may well have been more likely than not to be improvements.

[2] Perhaps the most important of these works are (Boyd and Richerson 1985; Boyd and Richerson 1996; Henrich and Boyd 1998; Boyd and Richerson 2000; Richerson and Boyd 2001b; Henrich and Boyd 2002;

<u>II Dual Inheritance and Cultural Evolution</u>

Gould and Lewontin, in <u>Spandrels</u>, argue that the practitioners of the neo-Darwinian synthesis exaggerated the free power of natural selection to craft adaptation and disparity. But they do not doubt that natural selection plays a central in the evolution of complex adaptation and they and others have begun to explore the features of biological inheritance that explain both the power and the limits of selection. In reviewing this work, I argued that three general properties of biological inheritance are central to the power of selection-like mechanisms to generate adaptation (Sterelny 2001; Sterelny forthcoming a).

First, inheritance systems must somehow <u>block outlaws</u>, since (as the group selection debate made clear) adaptive complexity at a level is often disrupted by the pursuit of individual interest one level down. Second, inheritance systems must allow the <u>stable transmission</u> of phenotypes over the generations. Evolutionary innovation depends on cumulative selection. If complex adaptations are to evolve, biological organization must be reliably rebuilt over many generations, not a few. It is for this reason that Richard Dawkins has emphasised the 'replicator condition' in defending gene selection (Dawkins 1982a). If a gene is mis-copied between, say, the F1 and the F2 generation, then the new version is faithfully transmitted to the F3 generation and beyond. Dawkins has argued that unless inheritance satisfies the replicator condition, cumulative selection and the gradual evolution of complex adaptive structure is not possible. For small improvements are not preserved as the basis for further cycles of variation and selection. However, though high fidelity replication may be necessary for the stable transmission of phenotypes across the generations, it is not sufficient. For the high fidelity replication of genes would not lead to the stable transmission of phenotype similarity if development were too sensitive to noise; if small variations in the developmental environment in which a gene functions typically made enormous differences to its phenotypic effect. For

---

Henrich and McElreath 2003; Richerson, Boyd et al. 2003; Gil-White forthcoming; Richerson and Boyd forthcoming).

then small variation in the set of genes transmitted across the generations would not reliably correspond to small variations in phenotype. The phenotypic effects of replicators must be stable too.

Third, selection depends on the generation of variation. That has a number of preconditions, one of which is developmental modularity. Unless development is to some important degree modular, selection will be unable to move a lineage away from its current organization (Wimsatt and Schank 1988; Wimsatt 2001). If the developmental program of an organism is richly interwoven so that the development of any given trait is connected to that of many others, then focal trait will not be able to change without that causing many other changes. Yet as the number of changes in a organization go up, so too does the probability that one of the changes will be a fitness catastrophe. The more modular the developmental network, the more contained are the consequences of change. Thus Wimsatt and Lewontin have pointed out that for adaptive change to be possible, developmental mechanisms must allow some traits to change without that affecting others {Wimsatt, 1980 #771); (Lewontin 1978). Unless that condition is satisfied, variation in the lineage will be highly constrained.

While there are ongoing controversies about the extent to which gene-based inheritance is an outlaw-protected, high-fidelity variation-generating system, there is no doubt that biological lineages often show significant evolutionary plasticity, and they do so because genes are replicated precisely and expressed in development environments that typically do not vary sharply from generation to generation. Given this conception of biological inheritance and of its evolutionary importance, dual inheritance systems are a natural first hypothesis in explaining the accumulation of cognitive capital. Dual inheritance theory models cultural evolution closely on genetic evolution, and cultural inheritance closely on biological inheritance. The central idea is that there are features of human psychology — most obviously imitation learning and language — and features of human social environments — most obviously long periods of juvenile dependence — that result in a high fidelity flow of ideas from parents to their offspring. These ideas — these cultural variants[3] — affect the behaviour and hence the fitness of these children. If a father transmits an improved stalking technique to his sons; if a mother transmits more reliable means of spotting the location of underground storage organs to her daughters, then these children will be more successful than others in their group, and they will transmit these cultural variants, in turn, to their own children. These culturally-based phenotypic traits will thus both increase in relative frequency and become available as a platform for further improvement.

Dual inheritance theory is obviously related to meme-based theories of cultural evolution, but on this picture it is the biological fitness of human agents that explains the spread or contraction of cultural variants: the fitness of those variants themselves plays no causal role in this picture[4]. I shall treat this dual inheritance picture as the base model of cultural inheritance, and in the next sections we shall see how this base model needs to be both extended and revised to explain the accumulation of cognitive capital. In the final section of the paper we will then consider some constraints on the power of cultural evolution, if that extended and revised model is anywhere near being right. The plan is as follows. I begin with a challenge, for the dual inheritance model seems to imply that the flow of information between the generations is memelike: accurate, discrete, vertical (i.e. from parents to their children). Yet there are persuasive reasons for thinking that this picture seriously mis-characterises human cultural learning. Boyd, Richerson and their co-workers respond to this challenge by attempting to show that the evolutionary accumulation of cognitive capital does not depend on a memelike flow of cultural variants across the generations. Their picture involves a shift from pure dual inheritance models of cultural evolution to models in which transmission between non-relatives (so-called "oblique transmission") is crucial.

---

[3] I will use this Boyd-Richerson term rather than "memes" to avoid prejudging the similarities between cultural and genetic evolution.

[4] It is also true that these dual inheritance models take cultural variants to be the information (and misinformation) transmitted between the generations, not the cultural products — artefacts, medicines, food — that are the products of the use of this information. This is not a forced choice, and a full-blown meme-theory might take artefacts themselves to be memes. For artefacts — like genes but not like ideas — are sometimes copied from templates.

In sections IV and V I discuss this response, and argue that it needs to be both supplemented and modified in important ways. The supplement: in my view the mechanisms on which Boyd and Richerson rely are effective only if cultural group selection is powerful. For I shall argue that the Boyd-Richerson models partially mischaracterize both the costs and the benefits of social learning. So the accumulation of culturally transmitted information depends not just on individual biological fitness (as in dual inheritance models) but also on the fitness of groups. The modification: I argue that while Boyd and Richerson can explain the preservation of the common knowledge of generation N in generation N+1, they do not explain how rare innovations can be preserved and amplified. So in section VI I argue that in explaining the evolutionary accumulation of cognitive capital, three transmission patterns are important, each with its own evolutionary causes and consequences. One is scaffolded and hence accurate vertical transmission from parents to children. Another is oblique but discrete transmission: transmission in which salient teachers in the N generation transmit their particular version of a skill to particular individuals in the N+1 generation. A third is oblique but diffuse transmission: where the N-generation as a whole transmits their lore to the next generation.

I conclude in VII with an analysis of the consequences of these ideas for cultural evolvability. In what circumstances, and with respect to what domains, can the mechanisms of cultural evolution build extensive information bases in support of adaptive actions? In short, this paper is a sketch of the nature and limits of cultural evolution.

III The Meme Paradox

I have suggested that just as complex biological adaptations are typically built by accumulating small improvements, so too most of the physical and informational tools that have enabled humans to survive have been built by the accumulation of small improvements. Perhaps some technologies are the result of a single individual breakthrough, as some able agent figures out a good trick and produces a near-optimal design all in one go. But natural history data bases; knowledge of the resources and dangers characteristic of a territory; much plant and animal husbandry; and most early technology, was built step by step. These tools are made by trial and error winnowing ideas and by keeping improvements. For most of human history, the informational resources that enabled us to lead human lives were accumulated gradually by something like a trial and error process.

Yet this conclusion leads to a problem. For the dual inheritance model of the trial-and-error accumulation of cognitive capital seems to presuppose that discrete units of information (and sometimes mis-information) are transmitted with high fidelity across the generations. On this picture, the cumulative evolution of adaptation depends on small, accidental improvements being preserved, to become the basis of further improvement. Intergenerational transmission is accurate, but when a new variant does occur, that variant (should it survive) is itself copied to the next generation. Dual inheritance theory takes the flow of cultural information to be like the flow of genetic information. Cultural variants — memes — should be like genes. And yet it seems that they are not.

There are two major challenges to the picture that models meme-flow on gene-flow. First, the very cognitive sophistication needed for accurate transmission of information across the generations destabilises parent/offspring similarities. For sucking information out of parental minds is not a dumb, blind process. The accurate transmission of information, especially about technique, across the generations seems to demand the capacity to imitate: to learn how to perform a skill by observing a model in the upstream generation. Imitation, however, turns out to be a very cognitively sophisticated capacity. Social learning is common, but only humans routinely learn new abilities by imitation (Tomasello 1999a; Tomasello 1999b; Whiten 2000; Heyes 2001). Yet these very abilities undermine the fidelity of the intergenerational flow of information. For they increase the extent to which an agent’s beliefs and behaviour are sensitive to his or her own experience, and they decrease the extent to which these beliefs and behaviours are the result of blindly copying those of the previous generation. Patterns of behavioural similarity will not be transmitted deeply through the generations.

Suppose, for example, that generation N has a stock of information about the appropriate techniques for growing corn. How reliably will this information flow to N+1? As Boyd and Richerson point out, this depends on the strength of "decision-making forces" at the N+1 generation (Richerson and Boyd forthcoming). To the extent that an N+1 agent is sensitive to his own experience of corn-growing and those of his neighbours and friends, the heritability of information from generation N will decline. And surely agents are sometimes sensitive to their own experience and to horizontal inputs: the "decision-making forces" of Boyd and Richerson are sometimes strong.

Moreover Sperber, Atran and Boyer have all argued both that the flow of information between the generations is not a template copying process and that this is in itself important (Sperber 1996; Boyer 1999; Boyer 2000; Atran 2001). Information in the mind of an agent of the N generation has to be manifested in that agent's action for it to flow to individuals in the N+1 generation, for silent memes die with their carrier. Moreover, that information must be inferentially reconstructed by the N+1 agent from the actions of the N agent. And this inference, they point out, will typically be difficult, error-prone and biased by the intrinsic structures of the human mind. The mechanism of cultural transmission makes it less likely to be accurate and reliable. That is particularly true if the inferential extraction of information from N-generation minds takes multiple trials. For then, as Sperber points out, the N+1 generation agent will often assemble his or her own version of the cultural variant from various sources. That variant will then be a <u>blend</u> of various N-generation versions.

This critique of meme-like models of cultural evolution consists of three elements, and it is important to distinguish them, for they have quite different consequences for dual inheritance models. The three ideas are: (i) cultural variants are not <u>copied</u>: information flow is inferential reconstruction, not copying; (ii) information transfer is typically not accurate; (iii) human minds have intrinsic characteristics that make some ideas salient and memorable, and others less so. I will take these in turn. First: Sperber points to a real difference between ideas and genes, for genes really are copied. Except those few arising newly by mutation, every gene in me has a distinct, identifiable ancestor in one or another of my parents. Genes form true lineages of ancestor-descendant pairs. But my ideas are not links in true lineages. Even if my ideas about Bradman are content-identical to those of my father, they did not arise from a single copying episode. They will have been constructed from multiple episodes of hearing about Bradman from my father (and others). My Bradman-conception and my father's Bradman-conception are not two links in a lineage. This is a very important objection to meme-theory proper, i.e. to a conception of memes that see them as genuine lineages, with fitness values to call their own, coevolving with humans. But it is not an objection to dual inheritance theory. Even if the flow of information from my father to me is not a copying process, and hence my Bradman-idea is not literally a descendant of his Bradman-idea, it could still generate stable, selectable phenotypic similarities that distinguish us from other members of our population.

The second idea — that the flow is inaccurate — is more threatening but less convincingly defended. Scott Atran, for example, describes a series of his own experiments, about what people understand about assorted political-religious slogans. Unsurprisingly he found wide variation in individual interpretation of these slogans (Atran 2001). These examples show nothing about the accuracy with which skills and other adaptively important information can be transmitted. In section VI, I shall resist the idea that culturally mediated information flow between individuals is typically too error-prone to support the evolutionary construction of adaptation.

The third idea about the intrinsic characteristics of the human mind is relevant and important. But these characteristics do not imply that there are no accurate inter-generational flows. Rather, they imply that not all information will flow with equal reliability. Information of the right type — salient and memorable — should be highly transmittable. In contrast, information of the wrong kind will be difficult to accumulate by transmission and incremental improvement. I will return to these issues in the final section when I discuss the kinds of information that can accumulate by cultural selection, and kinds that cannot. In the next two sections I shall explore a response to this challenge that rejects the idea that we need a meme-like theory of cultural inheritance to explain the accumulation of cognitive capital.

IV Cumulative Evolution Without Meme-like Replication?

I have just outlined an important challenge to dual inheritance theories of the accumulation of cognitive capital. One response to this challenge is to shift away from strictly vertical models of the culturally mediated flow of information. In human groups, information flows obliquely, from members of the N generation to unrelated members of the N+1 generation, and it flows horizontally, within a generation. Boyd, Richerson and their co-workers argue that these aspects of culturally-mediated information flow compensate for the fact that one-on-one linkages between parents and their offspring are often error-ridden. In the next two sections I shall argue that this response to the meme paradox is only partially successful. Oblique and horizontal flow explains how information that is the common property of the N generation — common knowledge — can be transmitted to the N+1 generation despite high error rates in one-on-one transmission. But these models do not explain how rare innovations in the N generation can be transmitted intact to the N+1 generation.

Richerson and Boyd accept that the transmission of cultural variants is importantly different from the transmission of genes. Biased transmission and guided variation operate to reduce the similarity between the set of cultural variants that characterise the behavioural repertoire of an agents' parents and those that characterise that agent. For an agent can non-randomly change the information to which she is exposed, and then transmit those changes (this is guided variation). Furthermore, agents can select their cultural parents; they can choose models from whom to copy. Finally, as the cognitive anthropologists note, the content of cultural variants can make them more or less likely to be copied (and if copied, copied accurately). These are forms of biased transmission. So they accept the Sperber-Atran-Boyer argument that the flow of cultural information is not meme-like. Even so, they think there is cumulative cultural evolution, and they have developed a series of models designed to show that mechanisms of cultural evolution can build adaptive capacities incrementally, despite low-fidelity inferentially-mediated information flow across the generations (Henrich and Boyd 2002). They undertake to show how incremental evolution is compatible with noisy, inaccurate and biased flow of cultural information across the generations. They argue that the processes of biased transmission can often correct these inaccuracies, resulting in population level flows of information that are much more reliable than one would expect, if one just focused on the flow of ideas from individual to individual (Richerson, Boyd et al. 2003; Richerson and Boyd forthcoming).

Think first of the importance of conformist biases: that is, a tendency to adopt the most common version of a local practice. Suppose that copying is seriously inaccurate: imagine an mimic has a 20% failure rate. That is, when a member of the model generation produces behaviour X on the basis of cultural variant A, 20% of the time the mimic takes the model to carry cultural variant B. If mimics learned only from their parents, such a failure rate would take you from fixation of A to a random mixture of A and B in only a few generations. Suppose, however, that mimic is exposed to five models and has a "majority-rules" learning bias. Under such circumstances, despite their high failure rate, almost all mimics will believe that most of their models carry A. Hence they will set their own value at A, and the frequency of A will stay very high, despite the inaccuracy of the individual "copying" process.

So a high error rate, if coupled with a conformist learning rule, is capable of preserving the informational resources of a group: the information embodied in A and common to generation N need not be eroded by error-prone transmission. For conformist learning adds redundancy: it is like sending the same message five times. However, the prospects for improvement under this regime — an adaptive shift from A to A* — would appear to be grim. For even if A* were to survive the 20% error rate, it would be very likely to be edited out by the conformist learning rule. The mimic would have to correctly identify the innovator as an A* bearer, and mistakenly identify two A-bearers as A*-bearers to set her own value at A*.

A somewhat similar moral emerges out of Boyd and Richerson's discussion of content-biases. Given that their carrier survives and reproduces, the flow of genes from parent to offspring is independent of the phenotypic effect of those genes. In most circumstances, meiosis is a fair lottery, and a gene in a parent has a 50-50 shot of making it to a fertilised egg, independently of the gene's effect on that egg. The same is not true of the cultural variants that a parent carries: some are more easily learned or remembered than others, and that biases the flow. Biases of this form (so-called "content-based biases") can support the accurate flow of cultural variants across generations, even if the actual learning mechanisms are noisy, for they can correct copying errors.

Boyd and Richerson discuss this mechanism through an extended discussion of the agricultural practice of share-cropping. In a share-cropping system a tenant works the land, and instead of paying cash pays a proportion of the crop as rent. It turns out that if agents were ideally rational, we would expect a wide range of farmer/owner divisions of the crop. For variations in soil productivity result in different divisions of the crop becoming the best stable bargain. Empirically, it turns out that where share-cropping is a standard land practice, there is only a few farmer/owner divisions of the crop (2-1, or 1-1) (Richerson and Boyd forthcoming). In Sperber's terminology, these are "basins of attraction". These crop division rules are salient and easily remembered, and so can be transmitted reliably and accurately between agents. In contrast, others are not, and will disappear. Basins of attraction that depend on inherent features of the mind turn cultural variants into something genuinely meme-like. A continuous array of possibilities — all the potential crop division rules — are transformed into a small number of discrete, reliably transmitted and selectable variants. The very salience and memorability of these simple rules makes selection important: it will determine the dominant practice. However, where such content biases play a crucial digitalising role, the range of variation will typically be quite restricted. Not many variants of a practice will be easy to remember and transmit.

So far, the models of Boyd, Richerson, Henrich and their co-workers seem to show that the difference between genes and cultural variants is compatible with the stable flow of information between generations. Once a capacity is established in a population, its informational basis can be transmitted reliably at a population level despite error-prone transmission in specific individual learning episodes. Redundancies and biases compensate for error. But these stabilisation processes depend on blending individual signals. So how could such mechanism generate the cumulative construction of new variants?

In their work with Henrich, they develop a model in response to this problem. The idea is that even when continuous traits are transmitted in an error-prone system, there can still be cumulative evolution if the selection effects are strong and if mimics preferentially copy the most successful agent (or agents) of the previous generation. The idea is that (a) all agents attempt to copy the N generation master of a particular skill/capacity. That mastery (the model assumes) is based on a particular continuous quantity, X. Higher values of X result in more skilful performance. No mimic will copy X perfectly, and indeed most of them will take from the master a worse version of X. But depending on the numbers of individuals trying to copy the N-generation master, some of the downstream generation will get lucky. Accidentally, they will have a higher X-value. In turn, the beneficiary of the lucky error will become the N+1 generation master, and will thus become the preferred model of the N+2 generation. Their skill will be based on X+n, which will become the new learning target. Selection can thus ramp up the X-average of the population, despite the lack of hi-fi copying between the generations (see Henrich & Boyd, 2002).

I have no doubt that the mathematics of this model works as advertised. But the mechanism is unconvincing. In their discussion of cultural selection reliably transmitting information at the population level without accurate replication at the individual level, Boyd and Richerson supported their formal models with plausible illustrative examples. In the case of cumulative evolution without replication all we get is the model, and that is no accident[5]. For it is hard to see a psychologically plausible analogue of the process they model. It is very strange to treat an information structure underlying a public capacity as a continuous quality. The manifestation of the capacity may well be a continuous quantity: hunting success; more durable pots; bows with more power. But the systems of information and skill underlying these capacities do not seem to be continuous quantities[6]. To the extent that imperfect copying is a problem, it is not because there must be some measurement error while trying to match a quantity. Rather, it is because the informational basis of skill is only partially manifest in any exercise of it. Skill consists in being able to respond effectively in a range of different circumstances, demands and materials. Unless each kayak is an exact copy made by identical means, a kayak-maker does not manifest all his skills in making any one kayak. So the model's implicit picture of the relationship between manifest capacity and its underlying informational basis is peculiar and implausible.

Moreover, as the size of a random change in an adaptive system increases, the chance of its being an improvement falls. A large blundering change in a mimic's reconstruction of the kayak-maker's craft will almost certainly degrade the skill. The real improvements made by luck will be small and that generates a covert tension in the model itself. It makes the mechanism of biased transmission less likely to work. For how is the N+1 master to be recognised as such? Other agents only see a small sample of his life. If the edge is small, they will be in no position to notice that Alphonse is a marginally better kayak-maker than Abu. Hence the very best techniques will not be differentially copied. In a noisy world "copy the best" is an accessible strategy only if the best is <u>markedly</u> better than the rest. That is especially true given that people reason poorly about statistical data (more on this in section VII). If most improvements are small, it will be hard to identify the best, especially if observers have only limited and sometimes inaccurate information about others' activities. All mimics can hope to do is copy one of

[5] Gil-White does provide an example, but in doing so abandons the distinction between a skill and its underlying informational basis: he discusses tennis serves, and trying to copy the best serve (Gil-White forthcoming).
[6] The features represented are often continuous quantities: how long a clay pot should be baked; how heavy an arrow head should be. But it by no means follows that the representation itself is a continuous quantity:

the good, and the Henrich-Boyd model does nothing to show that this learning rule will convert small, lucky overshoots of the model's technique into a platform for further improvement. Thus while I think the Boyd/Richerson response shows how individually error-prone transmission may be compatible with the preservation of accumulated information, I doubt that it can explain the accumulation process. The Boyd-Richerson explanation of the reliable transmission of common knowledge makes mysterious the establishment of beneficial imitation.

Let me summarise my reservations about the Boyd-Richerson account of cumulation. First, I doubt whether the models themselves are quite as powerful as they suppose: in particular, in explaining improvement. Second, I think the application conditions of these models are more restrictive than Boyd and Richerson suggest. For these models depend on cultural learning and in my view the Boyd-Richerson group somewhat mischaracterize the costs, benefits and nature of social learning. For them social learning is learning on the cheap. It tolerates a high error rate to avoid the high cost of directly learning about the environment by trial and error. In the next section I will develop and discuss this "expensive information" hypothesis. I shall argue that though Boyd and Richerson recognise the error costs of social learning, they underestimate those costs. Moreover, they underplay the costs of information donation to the <u>source</u> of that information, and hence to its recipients as well. As a consequence while they recognise the role of cultural group selection in human life, they do not emphasise its pivotal role in the evolution of extensive social learning.

Thus the cost of social learning to individuals is higher than Boyd and Richerson recognise. That in turn has two consequences. First, extensive social learning in human social life is not typically explained by a cost-driven flight from direct individual learning. <u>There has been no such flight</u>. Much human learning is hybrid learning. Often our adaptations for social learning do not operate to <u>replace</u> direct trail and error learning about the environment but to <u>supplement</u> it, by making it more reliable and by reducing

---

"about an hour" or "about half as heavy as the arrow itself" are discrete representations. And in these models it is the representation itself that must be copied.

its costs. Second, group selection plays a key role: the costs to individuals are partially compensated by benefits that accrue to human social groups rather than to individuals alone; moreover, there are features of social groups that reduce the costs of social learning (in particular, certain error costs) to individuals.

V The Expensive Information Hypothesis

Boyd and Richerson argue that the evolution of cultural cognitive adaptations is driven by information/reliability trade-offs: culturally acquired information is less reliable than information acquired by trial-and-error but it is also less expensive to acquire (Richerson and Boyd forthcoming). The cost of error can make the price of trial and error learning very expensive indeed: imagine sampling what turns out to be poisonous to see whether it is edible. But it is more reliable. For not only is cultural transmission error prone; in a fast changing world, accurately transmitted information can be out of date. The ideas of one's social group may have been overtaken by environmental change. The amount of cultural learning, its content, and its sources will depend on these cost-reliability trade-offs[7]. The Boyd-Richerson group suggest, though, that in many circumstances these trade-offs favour learning from others, including nonrelatives, rather than learning directly from the world. Oblique information flow is important in the Boyd-Richerson picture as part of their solution to low fidelity transmission; it makes cultural learning more reliable without making it more expensive.

---

[7] There are though, as these theorists note, extra complications here. For cultural learning can be more reliable if signals from the world are noisy. For:

> "suppose every individual is given a noisy signal from the environment about what is the best practice in the current circumstances. This information, for any one individual, might give them a 60% chance of noticing that blowguns bring back slightly larger returns than bows. Thus, using ordinary individual learning alone, individuals will adopt the more efficient hunting practice with probability .6. But if an individual samples the behaviour of 10 other individuals and simply adopts the majority behaviour, his chances of adopting the superior blowgun technology increase to 75% ((Henrich and McElreath 2003) p131)".

Hence conformist mechanisms can increase reliability by allowing "individuals to aggregate information over the behaviour of many individuals. Because these behaviours implicitly contain the effects of each individual's experience and learning efforts… ((Henrich and McElreath 2003) p130)). Of course, a problem here is that other individuals' practices may not be independent data points: the uniform vote for blowguns might be descended from a single episode of trial and error learning.

I am sceptical of this analysis of the adaptive salience of oblique transmission. I agree that oblique flow plays a central role in the accumulation of cognitive capital, but not because oblique flow is as cheap as vertical flow and somewhat more reliable as a consequence of the data-pooling effect adding redundancy. In brief, my response is as follows: cultural flow has costs to both model and mimic, and this shifts our picture of the selective mechanisms involved. The relative role of vertical and oblique transmission is tied to a dynamic between group and individual selection. To the extent that the accumulation of cognitive capital depends on oblique flow to overcome the noisiness of one-on-one learning links, it also depends on cultural group selection being sufficiently powerful both to damp-down individual selection for restricting information flow to one's own offspring, and to weed out maladaptive cultural variants. To put the idea the other way around, if the Boyd-Richerson group are right about the importance of oblique flow, an evolutionary condition on cumulative cultural evolution is powerful group selection.

This response depends critically on my view of the costs of social learning, so let me elaborate on these. Notice to begin with that cultural learning is not free for the model: for <u>the information-donor</u>. Often imitators impose direct costs. That is especially obvious in foraging, when company seriously reduces efficiency (companions, and especially incompetent ones, make noises at the wrong time, fall over things, attract unwanted attention). But even when imitation imposes no direct cost on the target of imitation, the <u>relative fitness</u> of model and mimic changes as the mimic acquires a skill that used to be the model's alone. Yet in many cases, the model must co-operate with the mimic for the mimic to acquire the model's skills and information. The model is often in a position to constrain the flow of information: to make crucial details more or less salient to those who are observing; to answer questions fully, badly, or not at all; to keep or to disclose tricks of the trade. In those cases where the model can control the flow of information, since imitation is not free for the model, models will be both motivated and able to extract a price for their information. They will not impose those costs on their offspring, but surely we would expect them to do so on others. Unless cultural group selection dominates individual selection, we will expect those with information to extract a price from those without it. Social learning will not be free, and perhaps not even cheap, for those acquiring information. If, then, the payoff for social learning is that it is cheap compared to direct individual learning, then these additional costs should shrink the pool of cases in which we would expect to see social learning. Yet it is ubiquitous.

Moreover, the Boyd-Richerson group are themselves committed to recognising the costs of learning for the ignorant. For Henrich and Gil-White have elaborated this idea into an explanation of a novel feature of human social life: hierarchy and benefit based on prestige. Their core idea is that (i) within a population, agents will vary in their skill level; (ii) agents who are highly skilled make the best models; (iii) agents have some control over the flow of the information that forms the basis of their skill; (iv) being the target of multiple mimics imposes costs on the model; (v) hence highly skilled models are both able to impose a price on their information and are highly motivated to impose a price; (vi) depending on the skill and the price, mimics will be motivated to pay; (vii) in many social circumstances, the form of this price is deference and prestige. Prestige is a fuel for social success an agent can transmute into a range of relevant, concrete returns (Henrich and Gil-White 2001). This picture is plausible. But it undermines the idea that culturally acquired information is typically cheap. It comes with a social price tag: deference towards the source of that information. Costs are relative, of course, but this line of argument would lead us to expect social learning only where individual learning is very expensive. Is this what we find?

To deference costs we must add error costs. If you download the characteristic belief set of generation N you will download their prejudices and superstitions as well as their genuine insights[8]. The Richerson-Boyd group acknowledge these costs, but I think they tend to understate them. For error costs arise not just from environmental change since the original trial and error episode and from corruption of the signal in cultural transmission. Not all skills, procedures, techniques are portable across agents and circumstances, and failures of portability impose costs that these theorists underplay. Portability can fail because the rate of return on a procedure is situation-specific. The social location of an agent may make a procedure successful for her but not for others. If (for example) she is high status, activities which are safe for her are risky for others, for she can rely on appropriate support and help. Portability may also fail because the rate of return depends on intrinsic capacities and skills of the agent: their strength, their physical or social skills; their informational resources.

Error costs generated by failures of portability will not be important for conformist learning biases. If a practice is common, its rate of return is probably not highly sensitive to individual circumstances. But they will be important to the costs and benefits of copying the successful, and as we saw above, that learning rule is very important to the Richerson-Boyd account of cumulative cultural evolution in the face of error-prone transmission. For the rate of return to the successful very often depends both on their intrinsic capacities and their social position: from their control of resources; from support and alliances; and from deference. These costs of cultural learning in turn constrain the power of evolution-like processes to incrementally build complex behavioural adaptations, by selecting for a more conservative approach towards learning from others. We do indeed sometimes learn by imitating the successful. But that learning tactic comes with a price, and partly for that reason, and partly because of the danger of error, it has to be judicious.

Portability adds to the cost differences between vertical and oblique information flow between the generation. Copying parental skills will be free: you will not be charged for the information and they have no reason to maliciously mislead you. That is not always true of information from other sources. Moreover, to the extent that the success of procedures depends on the characteristics of the individual using the procedure and their

[8] As Henrich demonstrates with a nice case study of central America farming practices: an accidental prejudice against barley and in favour of wheat was transmitted reliably across a couple of generations (Henrich 2002).

circumstances, other similarities between parents and their offspring increase the chances that what worked for the parents will work for their offspring.

The state of play so far. Central and adaptive features of human life appear to have been built by an evolution-like process. In particular, the informational and technological resources essential for human life appear to have been constructed incrementally, generation by generation. Prima facie, that seems to require the accurate transmission of information in a dual inheritance system. Both memes and genes flow from generation N to N+1. And yet there are persuasive reasons to think that the flow of cultural variants from N to N+1 is likely to be error-prone.

Boyd, Richerson and their co-workers attempt to resolve this apparent paradox by arguing that though cultural transmission is indeed error-prone, cumulative evolution is still possible. For one thing, it is not always error-prone. For both parents and their children share innate content biases, which render some cultural variants memorable and transmittable, and these shared biases stabilise some cross-generation flow of cultural variants, though at the same time reducing their potential to vary selectably. Moreover, biased transmission is important. If children sucked information only from their parents, ignoring all others of the N generation, often error would degrade cognitive capital to worthlessness. But information flows obliquely as well as vertically, and this allows conformist learning rules to add redundancy to the signals of the N generation, hence increasing the scope for preservation of cognitive capital despite noisy transmission. Moreover, success biases — the learning rule of copy-the-best — allows selection to drive incremental improvement even under conditions of inaccurate transmission.

In my view, the Boyd-Richerson line does not fully resolve the paradox of cultural accumulation. They understate the costs of conformist and success biases, and this reduces the range of cases in which these mechanisms can stabilise the flow of adaptive information. In view of the deference-costs and the error-costs of biased transmission, I am sceptical about the idea that social learning is operationally[9] cheap but somewhat less reliable than direct learning, and that it pays its way in environments which change fairly rapidly but not very rapidly. More seriously, I think the distinction between individual trial and error learning and social learning is overdrawn: much crucial learning is hybrid; socially enhanced, supported and protected exploration learning. Moreover, the model of incremental improvement in a noisy world depends on implausible conceptions of the skill/information relationship, and unrealistically optimistic assessments of the ease of identifying the best, whose capacities can then be copied.

VI Cultural Evolution, Hybrid Learning and Niche Construction

The evolution of culture theorists have responded to the memes paradox by arguing that biased transmission and allied processes can underwrite accurate between-generation information flow at the population level, despite the fact that individual learning episodes are noisy and error prone. I agree that oblique transmission is important to the accumulation of cognitive capital, but by itself oblique transmission is insufficient to explain accumulation. This explanation needs to be supplemented: individual learning experiences can be accurate at an individual level so long as they are scaffolded in appropriate ways. In particular, the distinction between social learning and trial and error learning is an oversimplification that really matters. Most human learning is hybrid learning, and its accuracy derives from this hybrid character.

In discussing this idea, three linked ideas will be important. The first concerns the pattern of transmission between the generations. Suppose in a foraging community we find in the N-generation a set of similar but still slightly varying techniques for (say) making baskets and other containers out of cane, and in the N+1 generation we find a similar set of techniques. At the population level, then, information is be transmitted with reasonable fidelity between the generations. Nevertheless, there are at least three different transmission patterns that can maintain this similarity, and these have different causes

---

[9] Only operationally cheap: Boyd and Richerson emphasise the fact that there are considerable costs in building the special cognitive equipment needed for efficient social learning.

and consequences. Transmission might be strictly vertical: from mother to daughter. Transmission might be oblique but discrete: there may be a few respected craftswomen of the N generation, each of which teaches (say) all her extended family. In such a case, transmission is oblique, but each individual in N+1 has a particular and identifiable model in N. Or transmission might be oblique but diffuse: each individual in N+1 is influenced by many individuals in N. The default expectation is that these patterns will vary from group to group; from skill to skill within a group and perhaps even for a given skill within a group. As I have already noted, these transmission patterns are linked to selective environments. In general, individual selection will favour vertical transmission, perhaps modified by an information market, whereas group selection will favour some form of oblique transmission. For it is more reliable. Transmission pattern and selective environment are in turn linked to learning mode. For there has been selection on groups to scaffold the acquisition of crucial skills. Like many other organisms we modify as well as react to our environment, and in our case those modifications include the collective construction of the developmental environment of the next generation. Much of human learning is hybrid learning, and that is often a consequence of selection on groups to socially enhance direct exploration learning.

Spectacular examples of hybrid learning can suggest that the mix of social transmission with individual exploration is powerful but rare. Heinrich mentions such an example in his discussion of the role of cultural transmission on agricultural practices:

> "Johnson (1971) describes how a Brazilian sharecropper, after observing a new method of planting bananas at a "technically advanced plantation" then performed a controlled experiment in which he planted rows of bananas with the new and traditional method — the traditionally planted rows acted as control group for comparison with the new method. Here the farmer first acquires an idea from a 'prestigious cultural model' and then experiments with the idea, before incorporating it into his behavioural repertoire. He did not arrive at the new method de novo, through calculation. He copied the prestigious model, then experimented. In my work with mapuche farmers I've found similar cases of prestige-based transmission and experiment" ((Henrich 2002) p34)

But while this is a vivid example, it is important to recall that all apprenticeship-style learning is socially mediated trial and error learning: “Is this right?” is a question that echoes down the ages, underlining the fact that cultural feedback on trials often supplements or replaces the rather more dangerous feedback from the world. For example apprentice craftsfolk can use exemplars to guide their actions in ways that support accurate skill acquisition. In some cases, artefacts can be used as actual templates in place of plans or blueprints. Somewhat more often (I suspect) they can be used as reference points at intermediate stages of construction, to check that all is going as it should be[10].

The use of artefacts as exemplars and guides illustrates a more general theme: humans have important biological adaptations for cultural learning: imitation, theory of mind and language. These do not operate in social worlds unchanged by their evolution. The existence of these capacities has initiated a cascade of changes whose effect is often to enhance the reliability of these very mechanisms. An important consequence is that there is an important range of cases in which one-to-one transmission is accurate and reliable. For example the accurate flow of information between the generations is scaffolded by the cultural evolution of explicit and implicit teaching, which in turn, as Raewyn Brockway has argued, depends on theory of mind, and especially perspective-taking abilities {Brockway, 2003 #1224}. It is much easier to teach when you understand what an agent does and does not know. Intergenerational flow is also scaffolded by epistemic technology: templates, well-designed specialist vocabularies, and mnemonics. Stories, cautionary tales and the like keep key episodes alive. More recently, it has been scaffolded by enduring public symbolic representations. Moreover, new techniques and discoveries induce changes in lifeways that in turn change the flow of information to the young in ways that make those new techniques easier to master. The transmission of some new techniques can become self-reinforcing. For example, as animal husbandry becomes increasingly important to a group, kids will grow up surrounded by animals and by people looking after them. As a result, they will have rich and repeated opportunities for socially guided trial and error learning about animals and how to care for them.

---

[10] Boyd and Richerson see this point, but only in passing. For they note that there could be selection for enhanced cultural learning if it made trial and error learning more efficient (Boyd and Richerson 1996)

My line of argument here connects to an important general theme in evolutionary biology: agents often respond to environments not just by adapting to them but also by changing them. Organisms partially construct their own environments. This was an important theme in Lewontin's earlier work (Lewontin 1982; Lewontin 1983; Lewontin 1985), and it has been developed further in that of Odling-Smee, Laland and Feldman on niche construction (Odling-Smee 1994; Laland and Odling-Smee 2000; Laland, Odling-Smee et al. 2000; Odling-Smee, Laland et al. 2003). They particularly emphasise the fact that niche construction often has downstream consequences. Agents often engineer their offspring's environment. Humans are downstream niche constructors par excellence. One important aspect of that niche construction is altering the epistemic environment of our offspring. We engineer the informational environment of our downstream generation, thus making for more accurate and reliable acquisition of key capacities.

Cultural group selection plays an important role in the evolution of hybrid learning. Downstream niche construction is not typically the action of a single agent. Parents contribute extensively to the learning environment of their children, but they rarely control those environments exclusively. In co-operative human environments, the learning environment of the offspring generation is in part the collective product of the group, especially as the children get older. So to the extent that cultural group selection has established co-operative norms, information will flow from the N-generation as a whole to the N+1 generation as a whole. Forager lifestyles depend on this flow of information, and these societies are famously egalitarian and co-operative (Boehm 1999). So suppose in a forager community:

> (a) There is the public broadcast of crucial information. Local natural history is stored and communicated through ceremonies, rituals, story-times, in which one or a few adults broadcast to audiences including the children of the group
>
> (b) Children learn crucial technical vocabulary from the community as a whole, and so they acquire linguistic cues of similarity and difference about their natural world diffusely, from their community as a whole.

> (c) In the upstream generation there is a division of informational labour, and salient individuals of generation N (elders, lore masters, respected hunters and craftsmen, shamans, etc) act as teachers, guides, mentors to many children of generation N+1.
>
> (d) The technology and techniques of many members of generation N are available as templates and props for the development of craft skills for most of generation N+1.
>
> (e) Children explore their environment with older, unrelated children and with other adults. So their learning environments is structured by the group as a whole rather than by their parents.

If cultural group selection has resulted in group-level practices of information pooling of this kind, then information flows collectively from generation N to generation N+1. In contrast, individual selection will favour vertical transfer, perhaps modified by a market in information. It will do so because vertical transfer preserves individual advantages within the group. Thus we might expect shamans; lore masters; those with detailed knowledge of the natural history of animals being hunted; those with complex technical skills (especially when the artefacts in question are difficult to reverse engineer: potters; canoe makers) to transmit their specialist knowledge selectively, to their sons and daughters.

There is some evidence that ritual knowledge really is transmitted vertically in most cultures. Moreover, there is archaeological evidence of local transmission of artefact-making technique over generations; technique including fine-grained stylistic features. So research in the evolutionary archaeology tradition shows that accurate transmission at the level of groups is possible, and some of this work suggests that family-sized units were capable of transmitting their own individual style across generations, thus when individual level selection generates strictly vertical transmission, it will tend to promote differentiation within a culture. There is good reason to believe that accurate vertical informational flow characterises some human social environments (Shennan 2002). Indeed, Henrich and Gil-White's model of the evolution of prestige presupposes this fact: there is no point in paying a deference fee unless one-on-one learning is accurate. There is no point in paying to learn from a high quality model if you cannot accurate copy the features of that model that conduce to excellence.

In contrast with individual selection for vertical transmission, to the extent that a co-operative milieu has been established and cultural group selection is strong, it will tend to favour shifts to oblique transfers of information from N to N+1. First, oblique transfer optimises individual learning: it gives the N+1 generation a chance to learn from the most skilled and knowledge members of the N generation, and it gives each member of the N+1 generation more relevant experience. If pot-making is a public rather than a private process, even if pot-makers do not vary significantly in individual skill, children will see pots being made more often. They have more learning opportunities.

Second, skills crucial to the group are less likely to be lost with oblique transfer. If pot-making or fire-starting is transmitted only in one or a few family lineages in the group, unlucky accidents can deprive the group as a whole of these skills. Shennan discusses the possibility that the inhabitants of the Torres islands lost their canoe-making capacities by such an unlucky accident [Shennan, 2002 #1280] pp 55-56). If not very many people possess a skill, the chance of losing it is surprisingly high. Tasmanian aboriginal people had very depauperate technologies by the time of first European contact, and archaeological evidence shows that technologies were lost <u>after</u> these peoples became isolated. These losses were very extensive: all bone-based technology; all fishing technology; the capacity to ignite fire; cold weather clothes; most stone technology. To fight and hunt, Tasmanians were reduced to one-piece spears, throwing clubs and rocks. Henrich's models suggest that the key variable was a fall in the absolute number of individuals with the crucial skill [Henrich, forthcoming #1330].

Indeed, Richerson, Boyd and Henrich suggest that this effect of group size on the reliability of transmission may explain the conservatism of hominid technology. Between the first appearance of stone tools in record two and a half million years ago and human technological takeoff one to two hundred thousand years ago, technological change seems remarkably slow. Perhaps this is an effect of small group size (Richerson, Boyd et al. 2003). Even with contemporary human capacities for cultural transmission, in small populations innovations are easily lost by unlucky chance. The takeoff of technology depended on the establishment of human social groups and population sizes that were large enough not just to afford specialisation (Ofek 2001) but to allow specialised traditions to endure in the face of unlucky accident. If these ideas are at all on the money, there will be strong selection on groups for some form of information pooling, for strictly vertical transmission of crucial skills will be very fragile.

Time to summarise the argument of this section. If, and to the extent that, individual level selection is responsible for cross-generational cultural transmission, and to the extent that hybrid learning and other canalisation mechanisms make this transmission accurate, then we will expect to see vertical and high fidelity transmission of information across the generations. This transmission pattern preserves innovations, so if they are advantageous they are selectable. Moreover, if those innovations are obviously successful, they are quite likely to spread horizontally to other members of the group through a prestige/deference informational market (and by information piracy). But from the perspective of the group, information held and transmitted strictly vertically is fragile and can easily be lost.

If a co-operative milieu is dominant, we might expect to see a more collective pattern of transmission from N to N+1. If that results in diffuse transmission, though information is preserved, it makes successful innovation less likely to be transmitted. If, a child's potting technique is a blend — an average — of the different potting techniques she has been exposed to, then a new and improved variation in N will not make it to N+1. If collective transmission makes a child's skills or knowledge base sensitive to all the exemplars he or she has been exposed to, then the conditions for the cumulative improvement of technique are undermined. However collective transmission may be oblique rather than diffuse; if, for example, key skills come accurately from a few salient individuals of the N generation. So long as these salient individuals — the key teachers of the N+1 generation — are innovators or sensitive to innovation, these mechanisms, too, can accumulate improvements. Cross-generation information flow is supported by niche construction. But the extent to which these flows are vertical, oblique or diffuse is obviously an difficult empirical issue, likely to vary between groups; over time and social structure within a group; and from skill to skill. My guess is that all these patterns are important, but their relative importance is very much an open question.

VII Cultural Evolvability

In the last three sections I have developed the idea that the informational resources needed for adaptive human action can be built incrementally, by evolution-like processes, through a combination of canalised and hence accurate vertical transmission and through niche construction, where the group as a whole collectively engineers the learning environment of the next generation. These mechanisms suffice for the conservation of the informational resources and, in favourable circumstances, their enhancement. I turn now to the issue of constraints. Where should we expect to find this process, and where not? Of course there is no question of an exhaustive survey of these issues; instead I shall develop a preliminary sketch of three of the factors bearing on these questions: (i) the selective environment; (ii) the architectural features of the human mind; (iii) the features of the target domain.

Consider, first, the selective regime. I have accepted the Boyd-Richerson argument that collective transmission plays an important role in stabilising the flow of information between the generations,, but have argued that this requires the establishment of a co-operative milieu, for donating information poses the same problem about the evolution of altruism as any other practice of providing resources for others.

The appeal to group selection is still very controversial in evolutionary biology proper, though it is not as heterodox as it once was. In contrast, cultural group selection is relatively unproblematic. Humans have psychological adaptations which enable them to police norms and enforce co-operation, and there is a significant body of work in experimental economics that shows both that people are willing to invest resources to punish defectors, and that other agents co-operate in response to this threat (Price, Cosmides et al. 2002). Boyd and Richerson call this “moralistic punishment” and have developed formal models to support their view that if small numbers of agents are prepared to punish to enforce norms, that changes the cost-benefit payoffs of other agents, and respect for those norms often then evolves. This mechanism does not intrinsically generate co-operative norms: moralistic punishment can drive arbitrary or even maladaptive norms to fixation (Boyd and Richerson 1992; Henrich and Boyd 2001). But selection at the level of groups will advantage those groups will prosocial norms (Sober and Wilson 1998). There is good reason to suppose that cultural group selection can be a powerful force in cultural evolution.

That said, it is also clear that there are important differences between cultures in the extent and nature of co-operation, and those differences will be relevant to the extent to which information flows obliquely in those social worlds. There has recently been a major cross-cultural study in experimental economics, probing the co-operative, prosocial dispositions in a range of small scale societies (Henrich, Boyd et al. forthcoming). In these experiments, members of these societies played ultimatum games for sums large enough to be meaningful to them (the local equivalent of two days pay). In an ultimatum game a resource is to be divided, and the first player proposes a division to the second player. That player can accept the proposal, in which case that division stands. Or the player can reject it, in which case both get nothing. Either way, the game then ends. Decision theory suggests that the first player should propose the maximally unequal division of the resource which offers the second player something, and the second player should accept that offer. Human players never conform to these predictions, and when these experiments are run with university students, they typically propose something like a 50-50 split. In this cross-cultural study, there was much greater variation: from hyper-fair offers through to offers around 80-20 splits. This study hints at the existence of considerable and continuing cultural diversity in norms of co-operation and sharing.

A full exploration of this diversity is vastly beyond the scope of this paper (for an initial sketch of some terrain see (Skyrms 1996)). But it is worth highlighting a couple of factors. Elinor Ostrom has highlighted some structural features of interaction which make co-operative interactions more likely and spirals to mutual defection less likely. She discusses co-operation in the context of farmers maintaining their water supply, but her picture generalises to other contexts. First, she suggests that symmetry between the agents helps, for it makes a fair allocation of cost easier, and likewise an assessment of benefit. As the tit-for-tat literature shows, a long time horizon and easy policing are important: trust is more likely to be built if it is easy to tell whether and what others have contributed. Information about past behaviour is obviously important, as is easy and face to face communication. This enables agents to make better choices about with whom to co-operate. Obviously, it will help if the contribution to the collective action problem not too expensive. Finally, it is important that agents have available graded sanctions for failing to keep agreements. Co-operative milieus are very unstable if any perceived defection causes the permanent retreat from any willingness to co-operate. In modelling co-operation, this is known as the “grim reaper” condition and it is often fatal to co-operation. Social situations characterised by these structural features are more likely to develop a co-operative milieu (Ostrom 1998). Hence cultures which typically conduce to such environments will be more co-operative, productive, successful. Perhaps one reason why forager societies are co-operative is that they do have these decision making environments.

Another factor is simply scale. Boyd and Richerson develop a good case for thinking that human prosocial emotions evolved in response to tribal scale social environments. This aspect of our evolved psychology generates a permanent tension between the demands of co-operation and co-ordination of large scale societies and our social instincts (Richerson and Boyd 2001a). To some extent, large scale societies can work around this problem through institutional designs which mimic smaller scale social life. But the tension is permanent and causes the permanent possibility of defecting factions rather than defecting individuals. Gangs, subcultures, cabals, guilds, cliques, juntas, and the like have strong in-group loyalty; they are part of larger social collectives, and they sacrifice the interests of those collectives in favour of the interests of a tribal faction within it. This behaviour is most obvious with economic resources, but we should find it as well with informational resources.

Let me turn now to architectural features of the human mind and their relevance. The cognitive anthropologists are certainly right in thinking that there are intrinsic constraints on the kind, quality, and accuracy of information that flows culturally from one generation to the next. Recent work in evolutionary psychology has over-emphasised innate constraints on human cognitive organization, and underestimated the effects of developmental environments on adult cognitive phenotypes (Sterelny 2003b). Nonetheless, there surely are important and developmentally entrenched features of the human mind that constrain the type of information that flows accurately between the generations. For example, we do not seem well-engineered for good statistical inference, for we are far too apt to see pseudo-patterns. In discussing the agricultural choices semi-subsistence farmers make when choosing which cereal to plant, Henrich illustrates the real bite of this constraint. He points out that these farmers face formidable inferential problem if they are to make an optimal choice. To choose between wheat and rival cereals, a farmer would have to recall both the average yield over a reasonable run of years and its variance. Moreover, he would have to be able to assess the significance of these results, taking into account potential confounding factors, and without (for example) overweighing the importance of a bumper crop in an especially benign year (Henrich 2002). He doubts that these farmers even attempt such calculations. Instead, they follow local custom.

This constraint is relevant to transmission mechanisms that depend on recognising success: either your own, or others. If human cultural transmission was uniformly vertical, and with no guided variation, human fallibility with respect to quantitative and statistical information would have no effect on the evolutionary fate of culturally-transmitted phenotypic traits. The bearers of some cultural variants would do a little better on average than others and they would faithfully transmit this slight edge to their descendants independently of anyone noticing that edge or understanding it. But to the extent that cultural evolution depends not just on success but on the recognition of that success, this constraint on our cognitive capacities is an important constraint on cultural evolution. Intrinsic constraints on human cognition are relevant not just to the accuracy with which a mimic copies or reconstructs the information held by a model, but to the choice of model(s). This consideration reinforces the more limited view of the importance of the “copy the best” transmission mechanisms I defended in sections III-IV.

We have just seen the importance of social regime to the mechanisms that accumulate and transmit the informational resources of a group. The mechanisms of both dual inheritance and niche construction are sensitive as well to intrinsic characteristics of the problem domain. Some domains of importance to us are more tractable than others. The development of humans weapons technologies has been far more uniformly impressive than the development of our health technologies. Until recently we have been far better at killing rather than curing. Perhaps this reflects historical contingency or differences in the strength of selection. Perhaps in an accident-prone and dangerous world (with their lowered life expectancies), being able to kill really mattered more than being able to cure. But equally likely, health is just more epistemically intractable. The causes of sickness are multi-causal. Likewise, even successfully interventions have probabilistic effects. Hence sophisticated statistical tools and experimental designs are often needed to tell whether a putative “cure” has worked. That is rarely a problem with weapons design[11]. Medicine scores low on “observability” — it is not easy to tell what works.

Moreover health environments are highly labile. Cognitive capital can accumulate by variation and the retention of successful experiments only if environments change relatively slowly. Cultural evolution may just be too slow to track health environments. Microbes themselves evolve very rapidly. Furthermore, the physical, biological, and demographic changes people make to their own worlds change the array and virulence of the microbes to which they will be subject. An increase in population size or density can ensure that a particular pathogen population will always have access to immunologically naive hosts and hence will not die out. New farming techniques can create opportunities for disease vectors to breed. Husbandry can create opportunities for pathogens to host-switch (Ewald 1994). Highly labile domains and those for which the success or failure of experiments is hard to judge will not support the cumulative growth of cognitive capital. Both of these problems infect human response to the challenge of pathogens.

There are also domains about which it is hard to run many experiments. Boyd and Richerson note ((Richerson and Boyd forthcoming), chapter 5) that many features of social organization have low trialability and low observability: you cannot try many alternatives, and it is hard to assess the success of those alternatives you do try. Consider the social arrangements that determine whether a son moves to his wife's village or vice versa. This is not a social arrangement an individual can trial in his own life: trying first the one, and then another. An experimental run involves an individual's whole reproductive life. Moreover, very often, an individual cannot probe this arrangement by assessing success or failure of others' family arrangements. Central features of family organization tend to be subject to norms which largely eliminate variability. If an agent lives in a family expecting his sisters to leave to live with their husbands, and his wife to come and live with him, then the chances are that all around him will observe similar customs. Even if he does live in a pluralistic environment with varying customs about family organization, these are likely to be highly correlated with other differences: in religion, ethnicity, and economic organization. This same confounding of marriage practices with other traits will also weaken the ability of cultural group selection to filter less adaptive marriage practices. For marriage rules will usually be part of complexes of social behaviour, and hence will not be individually exposed to cultural group selection. In contrast, physical tools and techniques for their use are often both highly observable and are easily trialed. You can give a new tool an trial and see whether it works, and you can observe the successes and failures of others. Hence we should expect to see, and do see[12], the more reliable accumulation of cognitive capital about technology, natural history, husbandry practices and the like but not family organization or in systems of religion and ritual knowledge.

---

[11] Of course, comparing relative effectiveness is a different matter: it may be very hard indeed to decide, all things considered, whether you should bomb or shell someone (or use a bow or a blow-gun whilst hunting). But relative effectiveness is an extra problem for medicine, too.

[12] Richerson and Boyd discuss examples of rapid cultural change in technique which satisfy just these conditions: the transition from atlatl-kayak hunting to firearm-umiak hunting, and the Aka transition to crossbow- hunting

This is, of course, the analogue of developmental constraints as they have been discussed in evolutionary biology. The existence of cultural constraints on the independent variability of such traits as these is no accident. For as we encounter one another, we face many co-ordination problems. In solving such problems, it helps to interact with other agents who have similar expectations about social interaction, family organization, child-rearing, norms of politeness and the like (Shennan 2002). Sets of such practices will tend to be come associated. They will both become a single package and as Gil-White in particular has argued, they will be advertised as such (Gil-White 2001). In contrast, technological and craft skills are more modular. As many folk in shared houses have found, you cannot live easily with someone with very different ideas about polite interaction; who prepares the food; what is a delicacy and what is disgusting; whether children should be treated gently or roughly. But there is no problem living with someone who prefers to use a different style of fishhook, or a different technique in weaving fish-nets. For the same reason, technologies often diffuse beyond the boundaries of the groups in which they are invented (Jordan and Shennan forthcoming).

Time to sum up. One of the most striking aspects of the ethnographic literature is the variability of human cognitive response to our environment. In some aspects, the people of a particular culture will respond to their world in an extraordinarily nuanced, subtle and informed way — and mostly in this paper I have concentrated on such examples. In other respects these very same people will seem barking mad: prisoners, for example, of vastly disabling beliefs about the polluting power of female menstruation. These variations may, of course, reflect historical contingency or the array of innate modules with which we are stocked (God neglected to update the menstruation-module). But perhaps, instead, it reflects both the power of cultural evolution to build culturally mediated cognitive adaptations, and the constraints on that power[13].

[13] Thanks to Peter Richerson, Peter Godfrey-Smith, Tim Lewens and Dennis Walsh for their comments on an earlier version of this paper.

References

Atran, S. (2001). "The Trouble With Memes: Inference versus Imitation in Cultural Creation." Human Nature **12**(4): 351-381.

Boehm, C. (1999). Hierarchy in the Forest. Cambridge, Mass, Harvard University Press.

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

Boyd, R. and P. Richerson (1992). "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups." Ethology and Sociobiology **13**: 171-195.

Boyd, R. and P. Richerson (1996). "Why Culture is Common but Cultural Evolution is Rare." Proceedings of the British Academy **88**: 77-93.

Boyd, R. and P. Richerson (2000). Memes: Universal Acid or a Better Mouse Trap? Darwinizing Culture: The Status of Memetics as a Science. R. Augner. Oxford, Oxford University Press.

Boyer, P. (1999). "Cognitive Tracks of Cultural Inheritance: How Evolved Intuitive Ontology Governs Cultural Transmission." American Anthropologist **100**(4): 876-889.

Boyer, P. (2000). Evolution of the Modern Mind and the Origins of Culture: Religious Concepts as the Limiting Case. Evolution and the Human Mind: Modularity, Language and Meta-cognition. P. Carruthers and A. Chamberlain. Cambridge, Cambridge University Press: 93-112.

Brockway, R. (forthcoming). Evolving to be Mentalists: The 'Mind-Reading Mums' Hypothesis. From Mating to Mentality: Evaluating Evolutionary Psychology. J. Fitness and K. Sterelny. Psychology Press, Hove.

Dawkins, R. (1982a). The Extended Phenotype. Oxford, Oxford University Press.

Ewald, P. W. (1994). Evolution of Infectious Disease. Oxford, Oxford University Press.

Gil-White, F. (2001). "Are Ethnic Groups Biological "Species" to the Human Brain? Essentialism in Our Cognition of Some Social Categories." Current Anthropology **42**: 515-554.

Gil-White, F. (forthcoming). Common Misunderstandings of Memes (and Genes). S. Hurley and N. Chater. Oxford, Oxford University Press.

Gould, S. J. (2002). The Structure of Evolutionary Theory. Cambridge, Harvard University Press.

Gould, S. J. and R. Lewontin (1978). "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." Proceedings of the Royal Society, London (Series B) **205**: 581-598.

Henrich, J. (2002). Decision-Making, Cultural Transmission and Adaptation in Economic Anthropology. Theory in Economic Anthropology. J. Ensminger, AltaMira Press: 251-295.

Henrich, J. and R. Boyd (1998). "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences." Evolution and Human Behavior **19**(2): 215-242.

Henrich, J. and R. Boyd (2001). "Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas." Journal of Theoretical Biology **208**: 79-89.

Henrich, J. and R. Boyd (2002). "On Modelling Cognition and Culture: Why Cultural Evolution Does Not Require Replication of Representations." Culture and Cognition **2**(87-112).
Henrich, J., R. Boyd, et al. (forthcoming). "Economic Man" in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-scale Societies. Co-operation, Punishment and Self-Interest. J. Henrich, R. Boyd, S. Bowles et al. Oxford, Oxford University Press.
Henrich, J. and F. Gil-White (2001). "The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing The Benefits of Cultural Transmission." Evolution and Human Behavior **22**: 165-196.
Henrich, J. and R. McElreath (2003). "The Evolution of Cultural Evolution." Evolutionary Anthropology **12**: 123-135.
Heyes, C. (2001). "Causes and Consequences of Imitation." Trends in Cognitive Science **5**(6): 253-261.
Jordan, P. and S. Shennan (forthcoming). "Cultural Transmission, Language and Basketry Traditions amongst the Californian Indians." Journal of Anthropological Archaeology.
Laland, K. N. and F. J. Odling-Smee (2000). Niche Construction, Ecological Inheritance and Cycles of Contingency. Cycles of Contingency. R. Gray, P. Griffiths and S. Oyama. Cambridge, MIT Press**:** 117-126.
Laland, K. N., J. Odling-Smee, et al. (2000). "Niche Construction, Biological Evolution and Cultural Change." Behavioral and Brain Sciences **23**: 131-175.
Lewontin, R. (1985). Adaptation. The Dialectical Biologist. R. Levins and R. Lewontin. Cambridge, Harvard University Press.
Lewontin, R. C. (1978). "Adaptation." Scientific American **239**: 156-169.
Lewontin, R. C. (1982). Organism and Environment. Learning, Development and Culture. H. C. Plotkin. New York, Wiley**:** 151-170.
Lewontin, R. C. (1983). "The Organism as the Subject and Object of Evolution." Scientia **118**: 65-82.
Odling-Smee, F. J. (1994). Niche Construction, Evolution and Culture. Companion Encyclopedia of Anthropology. T. Ingold. London, Routledge**:** 162-196.
Odling-Smee, J., K. Laland, et al. (2003). Niche Construction: The Neglected Process in Evolution. Princeton, Princeton University Press.
Ofek, H. (2001). Second Nature: Economic Origins of Human Evolution. Cambridge, Cambridge University Press.
Ostrom, E. (1998). "A Behavioral Approach to the Rational Choice Theory of Collective Action." American Political Science Review **92**(1): 1-22.
Price, M., L. Cosmides, et al. (2002). "Punitive Sentiment as an Anti-Free Rider Psychological Device." Evolution and Human Behavior **23**(3): 203-231.
Richerson, P. and R. Boyd (2001a). The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis. Evolution and The Capacity for Commitment. R. Nesse. New York, Russell Sage Foundation**:** 186-220.
Richerson, P., R. Boyd, et al. (2003). Cultural Evolution of Human Co-operation. Genetic and Cultural Evolution of Cooperation. P. Hammerstein. Cambridge, MIT Press**:** 373-404.

Richerson, P. J. and R. Boyd (2001b). "Built For Speed, Not For Comfort." History and Philosophy of the Life Sciences **23**: 423-463.
Richerson, P. J. and R. Boyd (forthcoming). The Nature of Cultures.
Shennan, S. (2002). Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution. London, Thames and Hudson.
Skyrms, B. (1996). The Evolution of the Social Contract. Cambridge, Cambridge University Press.
Sober, E. and D. S. Wilson (1998). Unto Others: The Evolution and Psychology of Unselfish Behavior. Cambridge, Harvard University Press.
Sperber, D. (1996). Explaining Culture: A Naturalistic Approach. Oxford, Blackwell.
Sterelny, K. (2000). "Development, Evolution and Adaptation." Philosophy of Science (supplementary volume) **67**(Supplement): S369-S387.
Sterelny, K. (2001). The Evolution of Agency and Other Essays. Cambridge, Cambridge University Press.
Sterelny, K. (2003a). "Last Will and Testament: Steven Jay Gould's The Structure of Evolutionary Theory." Philosophy of Science **70**(April): 255-263.
Sterelny, K. (2003b). Thought in a Hostile World. New York, Blackwell.
Sterelny, K. (forthcoming a). Symbiosis, Evolvability and Modularity. Modularity in Development and Evolution. G. Schlosser and G. Wagner. Chicago, University of Chicago Press.
Tomasello, M. (1999a). The Cultural Origins of Human Cognition. Cambridge, Harvard University Press.
Tomasello, M. (1999b). "The Human Adaptation for Culture." Annual Review of Anthropology **28**: 509-529.
Whiten, A. (2000). "Primate Culture and Social Learning." Cognitive Science **24**(3): 477-508.
Wimsatt, W. C. (2001). Generative Entrenchment and the Developmental Systems Approach to Evolutionary Processes. Cycles of Contingency. R. Gray, P. Griffiths and S. Oyama. Cambridge, MIT Press**:** 219-238.
Wimsatt, W. C. and J. C. Schank (1988). Two Constraints on the Evolution of Complex Adaptations and the Means of Their Avoidance. Evolutionary Progress. M. H. Nitecki. Chicago, University of Chicago Press**:** 231-275.