Fatal Attraction? Why Sperber’s Attractors do not Prevent Cumulative Cultural Evolution

Catherine Driscoll

ABSTRACT

In order to explain why cultural traits remain stable despite the error-proneness of social learning, Dan Sperber has proposed that human psychology and ecology lead to cultural traits being transformed in the direction of attractors. This means that simple-minded Darwinian models of cultural evolution are not appropriate. Some scientists and philosophers have been concerned that Sperber’s notion of attractors might show more than this, that attractors destroy subtle cultural variation and prevent adaptive cultural evolutionary processes from occurring. I show that Sperber’s view does not have this consequence—that even if there are attractors, cumulative cultural evolution can still occur.

1 Introduction

2 Why Attractors are Supposed to Pose a Problem for Cumulative Cultural Evolution

3 Why Attractors do not Prevent Cumulative Cultural Evolution

3.1 Reconstruction probably retains adaptively important features of cultural traits

3.2 Cultural selection may counteract the affects of attractors when cultural variation is adaptively important

3.3 Attractors may be numerous

3.4 Factors of attraction may often be working on the components of cultural traits

3.5 Attractors are highly variable, even within local populations

4 Conclusion
1 Introduction

In recent years there has been some significant debate between two groups of evolutionary social scientists working on the evolution of culture: on the one hand, the dual inheritance theorists (DITs) (Cavalli-Sforza and Feldman [1981]; Boyd and Richerson [1985]; Feldman and Laland [1996]; Durham [1991]) and on the other a group of evolutionary psychologists or cognitive psychologists and anthropologists influenced by evolutionary psychology who hold what is sometimes referred to as the ‘epidemiology of representations’ (EoR) view (Sperber [1985], [1994], [1996], [2000], [2006]; Boyer [1999], [2001]; Atran [2001]; Sperber and Claidiere [2006], [2007]; Claidiere and Sperber [2007]). Like some other debates in the sciences, this one has at times generated more heat than light; in reality, both groups are converging on a similar view about the nature of cultural transmission and evolution, and the differences between them are largely those of emphasis. Dual inheritance theorists emphasize the role and power of population-level cultural evolutionary processes, especially adaptive and cumulative cultural processes, in explaining cultural change. Cumulative cultural traditions’ being both cumulative and adaptive is important because it explains how the capacity for complex culture came to be selected for in humans (Boyd and Richerson [1996]; Richerson and Boyd [2000]). For the EoR group, especially Sperber and his collaborators, the emphasis has been in understanding the causes of stability in cultural traits, rather than cultural change, and the role that various forces, most especially the human species-typical psychology, play in permitting that stability. Sperber and his collaborators argue that cultural transmission processes are not simple mindedly Darwinian, in that they cannot be modeled in a psychologically shallow way, as ‘memes’ ‘replicating’ themselves; individuals have to reconstruct their teacher’s representations rather than copying them, and in the process will often transform the traits they are copying. How the trait is transformed during reconstruction depends greatly on many factors besides the content of the learnt representation, in particular, the nature and limits of the learner’s psychology, and ecological and social constraints, among other things. Because human cultural populations tend to share some of these factors—which Sperber calls ‘factors of attraction’—cultural traits tend to morph into a constrained range of types—attractors—as they are transmitted through those populations. Sperber believes that attractors explain the relative stability of cultural evolutionary processes despite the imprecision of cultural transmission.

Unfortunately, the debate has been marked by a number of misunderstandings. Sperber’s argument was initially directed at the dual inheritance theorists, probably because he mistakenly took their position as being a simple mindedly Darwinian ‘meme’ theory, rather than one that takes into account
the psychological forces that drive cultural evolution. Consequently, Sperber’s discussion of reconstruction and attractors has been taken as an argument against dual inheritance theory or, more precisely, against the possibility of adaptive cumulative cultural evolution. Various scientists working in dual inheritance theory (e.g. Henrich and Boyd [2002]; Henrich and McElreath [2003]; Richerson and Boyd [2005]; Henrich et al. [2008])—and possibly some philosophers (e.g. Sterelny [2006], [2007])—have taken the existence of attractors to mean that cultural variation necessary for cumulative cultural evolution is usually and inevitably destroyed, either at once or over the course of cultural transmission processes. This concern is not entirely unwarranted, because attractors do pose a *prima facie* problem for dual inheritance theory; however, I believe that Sperber’s attractors are essentially compatible with adaptive, cumulative cultural evolution; in this article I will describe a number of reasons why this is. I will proceed as follows. In the second section, I will explain in a bit more detail why cumulative cultural evolution is very important to current explanations of the evolution of human culture, how attractors are supposed by Sperber to work, and exactly why some scientists are concerned that they create a problem for dual inheritance theory. I will then, in Section 3, go on to offer five reasons why attractors, properly understood, will not prevent cumulative cultural evolution.

2 Why Attractors are Supposed to Pose a Problem for Cumulative Cultural Evolution

Dual inheritance theory purports to explain the range and variety of human behavior seen across cultures\(^1\) in terms of differential and cumulative cultural transmission processes. The reason why human beings engage in such a wide range of different types of behavior and produce such an extraordinary range of complex artifacts is because the ideas and representations responsible for those behaviors and artifacts have undergone processes of transmission and repeated refinement in human populations until they become highly adaptive and elaborated. Such processes occur, first, where there is relatively accurate copying of traits between individuals; second, where there are individuals occasionally generating new, more adaptive variants of cultural traits; and third, where humans have a cultural learning psychology composed of heuristics that permit them to acquire cultural traits that are more adaptive than average more often than not. DITs refer to such processes as *cumulative* cultural evolution; the adaptive power of cumulative cultural evolution is so important to the DIT’s because it is their explanation for why capacities for social learning

---

\(^1\) Dual inheritance theorists are also, of course, interested in how transmission processes of cultural traits affect the transmission of genes/phenotypic traits and vice versa (hence ‘dual’ inheritance theory).
are adaptively advantageous (Boyd and Richerson [1985], [1996]; Richerson and Boyd [2005], [2000]) and hence why they were selected for. The main benefit of social learning is that it avoids the costs of individual learning (i.e., learning by oneself how to behave adaptively in an environment)—the time, effort, and danger. Social learners, however, pay the cost of not having behavior learnt immediately from the environment—because they learn from others, who may themselves have learnt from others and so on, it is possible that the behavior they acquire is no longer adaptive, especially if the environment changes in a relevant way (Rogers [1988]). This means that, if social learning is to become widespread, there must always be some individual learners in the population to acquire newly adaptive traits for the social learners to copy. However, where cultural transmission is adaptive and cumulative, social learners have a significant advantage over individual learners in that they can potentially acquire traits much more adaptive than any individual learner could learn by themselves. So it is the possibility of cumulative evolution which is supposed to then drive the necessary selection for cultural learning psychology in humans. DITs accept that human psychology is responsible for cultural variation and for the selective transmission of cultural traits that permits adaptive processes; nevertheless, since cultural transmission is a phenomenon that happens at the population level, the nature, patterns, and distribution of cultural variation are partly or largely determined by the cultural evolutionary process and not simply (or even largely) by the psychology that generates and acquires it. This means that this distribution can be properly understood using the types of general and simple models of cultural processes the DITs use in their work (Cavalli-Sforza and Feldman [1981]; Boyd and Richerson [1985]; Feldman and Laland [1996]). Furthermore, if cumulative cultural processes are to be powerful and adaptive in the way the DITs need, it is important that a wide range of relatively fine-grained cultural variation be available and maintained in the population. DITs believe that human psychology is such that the cultural variation it generates is relatively unconstrained or, at least, sufficiently unconstrained that humans can regularly generate and improve on already existing cultural variation (Boyd and Richerson [1996]).

The central point of disagreement between Sperber and the DIT’s is over the relative accuracy of cultural transmission. The DITs believe that cultural

---

2 This is by no means the whole story. According to Richerson and Boyd ([2000]; Boyd and Richerson [1996]), if cumulative culture were always adaptive, it would be common in nature. It isn’t, probably because cumulative cultural evolution cannot occur without many social learners in a population, and social learning without cumulative culture is only adaptive under a restricted range of conditions. However, social learning may be better than individual learning in cases where the environment is changing, but not too often, as it was in the Pleistocene when humans evolved. Cumulative cultural evolution probably also requires other psychological capacities which happened to be present in humans at that time due to other causes.
transmission must be relatively accurate if there are to be stable cultural processes: too much inaccurate copying, and cultural variation is lost. They believe, however, that cultural transmission processes are accurate (or at least, as I shall discuss shortly, accurate enough that the adaptively relevant features of cultural traits get transmitted). Sperber disagrees: according to him, people rarely copy other individuals accurately—indeed, they don’t really copy each other at all. Sperber ([2000]) uses an example from Richard Dawkins ([1999]) of schoolboys learning to fold an origami junk. Learning to fold the junk requires each schoolboy to watch another boy go through the various individual steps, and then to reproduce those steps himself in the correct order to produce the junk. However, in practice, learner boys do not actually reproduce what the teacher boy does. For example, one of the steps might be to fold the paper in half. If the teacher boy is not completely accurate—say the paper does not meet completely in the middle—the learner boys do not reproduce the actual behavior of the teacher boy—instead, they fold the paper exactly in half (or as near as they can manage). Furthermore, cultural traits are often acquired and synthesized from the behavior of many teachers (Sperber [1996]). A boy might watch several boys fold the origami junk before he forms his own mental representation of how it should be done; an individual might reproduce a piece of music that is a combination of the styles of several different original versions; a philosophical idea might be a combination or refined synthesis of the ideas of many other philosophers. In such cases individuals are not really copying any particular trait either.

The failure to copy, however, does not seem to result in a lot of cultural variation between individuals in populations—instead, cultural traits seem to remain quite stable. Sperber’s central question is how this stability is maintained. His explanation has two parts. First, cultural transmission is not so much copying as a process of reconstruction and transformation—that genuine copying of cultural traits ‘should be viewed as just a limiting case of maximal resemblance’ (Sperber [1996], p. 83). What the learner boys above are doing is not copying the teacher boy’s behavior but instead trying to copy his mental representation of the cultural trait—what he really intended to do. But the learner boys can’t really copy the teacher’s mental representations, because such representations are not immediately observable; instead, those representations have to be reconstructed—the learner boys have to attempt to guess what the teacher boy intended from clues such as his behavior, the context, their background knowledge, and so forth. In this case the boys reconstructed the teacher boy’s intention as ‘fold the paper exactly into the middle’ because of what they knew about people’s error proneness, their background knowledge about origami, and so forth.

The second part of the explanation for cultural stability is, according to Sperber ([2000]) that there are psychological, social, and ecological causes for
the way humans interpret their teachers and hence reconstruct cultural traits in the way they do (‘factors of attraction’)—for example, the clues used by the learner boys described in the case above. In many populations, many individuals share these factors of attraction. Where such factors are shared, new cultural variation introduced into populations tends, as it is transmitted from individual to individual, to degrade into a common form determined by these factors of attraction. These forms Sperber ([1996]) calls ‘attractors’. The term comes from a particular way of representing the cultural evolutionary process. Imagine the range of possible cultural traits as a space in which each point (or sector) in the space represents a possible cultural trait. The idea is that cultural innovations begin at some point or sector in the space. When the trait is transmitted to another individual and transformed it moves away from the point in the space it was at in the previous generation. Because the factors determining transformation during reconstruction are shared and because individuals in those populations tend to transform cultural traits in the same kinds of ways, there are points in the space that cultural traits are more likely to move towards each time they are transmitted (of course, the tendency to be transformed in this way is only probabilistic). As a consequence, cultural traits tend to be ‘drawn towards’ these points with each replication, as if they were black holes sucking in all the cultural variation—hence the name ‘attractors.’ In the origami case, above, the attractor is ‘fold the paper exactly into the middle’ which attracts all the cultural variants involving slightly skewed folds. This is what accounts for cultural stability—variations from that cultural norm tend, over time, to be removed from the population during transmission from person to person.

Sperber’s view, however, has provoked some controversy because it appears *prima facie* to pose a significant problem for the adaptive cumulative cultural evolution which is so important to the dual inheritance theorists (see e.g. Henrich and Boyd [2002]; Henrich and McElreath [2003]; Richerson and Boyd [2005]; Henrich et al. [2008])—and maybe (Sterelny [2006], [2007]). Cumulative cultural traditions require not only that variation be preserved in populations, but that new variation, especially adaptive variation, also be able to be preserved when it arises, and incorporated into cultural traditions. If there are attractors, then there is a force operating in populations that not only can corrupt cultural variation in *individual* cases, but does so in a persistent and *directional* fashion. The dual inheritance theorists have argued that mere noise in transmission won’t prevent cultural evolution—it can even help it if individuals are using learning biases, because this noise can create new, more adaptive variation for learners to choose from (Henrich et al. [2008]). But persistent, directional destruction of variation by attractors does not help, since any new adaptive variation hit on during transformations due to attractors will itself be destroyed as the trait moves to the attractor. If new cultural
variation can be readily lost, then there can be no cumulative cultural processes, and no significant selective force to explain the origin of human adaptations for cultural learning. Furthermore, if attractors do readily draw in cultural variation it also appears that cultural evolutionary processes do not really explain the nature and distribution of cultural traits in populations, as opposed to the ‘factors of attraction’ that determine what the attractors are in those populations. Despite these legitimate concerns, I believe that attractors, at least as Sperber imagines them, will not necessarily prevent adaptive cumulative cultural evolution from occurring: significant cumulative cultural evolution is compatible with the presence of even quite strong attractors. Sperber’s view of cultural transmission is, therefore, largely compatible with the dual inheritance theorists’ story about the evolution of culture. In the next section, I will discuss a variety of reasons why this is.

3 Why Attractors do not Prevent Cumulative Cultural Evolution

In this section, I am going to discuss five important reasons why attractors will not necessarily—nor are even particularly likely to—prevent adaptive cumulative cultural evolution. As I explained in the previous section, cumulative cultural evolution is necessary to explain how humans have generated such an impressive array of complex artifacts and ideas, and perhaps more importantly, how humans come to have culture in the first place. If the dual inheritance theorists are right, cultural processes need to be powerfully and cumulatively adaptive to explain the origin of human cultural learning psychology: if Sperber is right about attractors, then attractors must not be able to prevent such adaptive cultural processes from occurring.

3.1 Reconstruction probably retains adaptively important features of cultural traits

The first problem that I want to deal with is the problem of the inaccurate copying of cultural traits. Sperber and his colleagues claim that cultural transmission is error prone; indeed, that cultural transmission makes the ‘copy’ of the trait deviate so much from the original that it perhaps should not be thought of as a ‘copy’ at all. Instead learners reconstruct cultural traits based on their own interpretations and expectations about what their teacher means—a reconstruction often transformed relative to the original variant. This is, \textit{prima facie}, a significant problem for the dual inheritance theorists’ view of cultural evolution as a cumulative, adaptive process because such a process requires that new, adaptive cultural variation persists in the population and accumulates during cultural processes. Too much inaccuracy leads to the corruption of this variation and means that cultural processes cannot
accumulate adaptively beneficial cultural variants in the population. Part of the problem with this debate is that Sperber and his colleagues and the dual inheritance theorists are not clear about what counts as ‘a lot of’ or ‘important’ error in cultural traits, and this lack of clarity may be exacerbated by the two sides focusing on different sorts of cases when trying to understand cultural transmission. Sperber and his colleagues tend to focus on cases where the trait transmitted is a story, a song or some other apparently adaptively neutral trait; the dual inheritance theorists concentrate on the transmission of traits that have clear adaptive value, such as tools, agricultural techniques, ecological information, and so forth. Since there is no clearly appropriate standard of accuracy for adaptively neutral traits, any error could appear significant; however, with adaptive traits, there is a standard (if a loose one)—even what looks like a lot of error might be insignificant if it corrupts none of the functionally and adaptively important features of the trait. Let me explain what I mean. If transformation during cultural transmission generally leads to individuals retaining the functional and adaptively important features of cultural traits, then adaptive, cumulative cultural evolution can occur. For example, a learner may not have the same exact mental representation as their teacher with regard to a pot she is trying to learn to make, nor preserve all of its detailed features; but if the learner can acquire the functionally relevant aspects of the representation (e.g. enough such that the pot appropriately holds water without leaking; has a spout that permits accurate pouring, etc.), then the loss of other details is irrelevant. The adaptive value of the pot as a way of storing water is retained (and may be improved over time).

What is needed to show that cumulative cultural evolution can occur, then, is some reason to think that reconstruction of cultural traits does generally preserve their adaptive features. Does the error that Sperber thinks occurs in cultural transmission amount to error in these features? Sperber’s belief that there is significant error in reconstruction stems from his concern that reconstruction is difficult for the learner. Sperber himself has done some important work in showing that understanding communication is a non-negligible task, and requires individuals to make an inference from the utterance and the context to what the communicator means. Individuals choose interpretations that first, would cause most changes in their beliefs (since the statement would thereby be informative), and second, that require the minimal amount of cognitive effort (since other individuals are usually not thinking in difficult or obscure ways when they are trying to communicate) (Sperber et al. [1995]). However, in this section I want to identify some forces that might be acting in the learner’s favor to make manageable reasonably accurate learning of at least the functionally and adaptively relevant features of cultural traits. So let’s first have a look at why Sperber and his colleagues argue that
Fatal Attraction

reconstruction tends to lead to the transformation of cultural traits. They identify two main factors—a newly acquired cultural trait...

...is likely to depart from the variants on which it is based both because some information may be lost in the process, and because the goal of acquisition is generally to acquire not a replica of other people’s variants, but, rather, a piece of knowledge or a skill that suits the individual’s own dispositions and preferences (Claidiere and Sperber [2007], pp. 91–2).

In other words, cultural traits can fail to be copies due to general error, or due to deliberate changes on the part of the learner due to their own unique dispositions and desires. Let’s look at deliberate changes first. Deliberate changes are supposed by Sperber to be due to the learner having unique interests in the trait in question, and what she wants to do with it. However, at least with regard to those traits that are most adaptively relevant (and on which the dual inheritance theorists’ explanation for the origin of cultural capacities hangs), it is not obvious that the dispositions and preferences of learners are likely to lead them to acquire traits that are different from their teachers in those regards most functionally and adaptively important. I think Sperber is right that learners want to acquire cultural traits from their teachers with features which will be specifically useful to them, and that what they want to pick out will depend on their desires and situation. However, there’s no reason to think that in the case of the most adaptively relevant traits (and especially with regard to their most adaptively relevant features) people do have very different desires than their teachers, because learners and teachers share much the same biological goals and drives, and are often sharing more specific ecological and social concerns as well. A learner wants to acquire the technique for building a fishing kayak because she knows that her teacher is interested in acquiring fish to eat, an abundant and useful supply of food in their environment; and she shares the need for an abundant food source of the same kind. When learning an adaptively neutral trait like a song or story, it is easier to see how individual preferences can change the acquired trait—the learner likes some parts of the story more than others and is more likely to remember and transmit them, alter the story to emphasize them, and so forth. It is less likely that this will happen with the kayak, at least insofar as the adaptively relevant features are concerned—while the learner might paint it a different color, she is unlikely to decide that she prefers it with holes poked in the bottom or without oil in the seams. The learner needs and intends to retain those features that make the kayak do what will be useful in helping her meet her biological and or social goals; of course, determining which features those are may be a difficult task, but given the teacher can make this clear, it is highly unlikely the learner will reject them in her own reconstruction. This suggests that in those cases most important for cumulative cultural evolution, Sperber’s second type of
transformation, deliberate changes from the original trait, will generally only
take the form of changes that are adaptively neutral or positive for the learner,
and neither will prevent cumulative cultural evolution from occurring.

What about general error, however? Even here, I think there are significant
forces acting in the learner’s favor. First, if all people have to reconstruct is the
most functionally important elements of the teacher’s trait, the difficulty of the
exercise is significantly reduced. The learner does not have to determine every-
thing that the teacher has in mind, only those features that are most significant
given their common adaptively relevant interests (and it doesn’t matter if she
also acquires any adaptively neutral elements—the only elements that need to
be right are the functionally and adaptively relevant ones). Second, the shared
biological interests of learners and teachers in the case of adaptive traits mean
that it is likely that the things the teacher will consider most important about a
trait or artifact and be trying to emphasize will be the things the learner will be
trying to learn (and need to learn if cultural evolution is to be cumulative and
adaptive. Furthermore, (and this is a point on which I think Sperber would
agree) in such cases, common factors of attraction can actually help learners to
acquire their teacher’s cultural traits relatively accurately. Because of their
shared species typical psychology, and because teachers and learners often
occupy common social and ecological environments, the trait the teacher her-
self has acquired and is teaching to the learner is likely to be the same trait the
learner will reconstruct; and their expectations about what is relevant are
likely to be similar. Hence, the interpretations of the teacher that come
most naturally to learners will tend to be what the teacher actually has in
mind. Hence, minimizing cognitive effort probably works in the reconstruc-
tion of the adaptive features of cultural traits.³

The biggest problem that remains is that not all cases of reconstruction
involve the teacher trying to communicate things that fit in with the learner’s
expectations—sometimes the adaptively important features they are trying to
teach may be things that are subtle or unexpected (that require, as Sperber
would say, a lot of ‘cognitive effort’). In such cases the learner’s expectations
might lead her astray. However, teaching situations are not usually one shot
interactions. Sperber ([2000]) discusses a thought experiment that shows how
expectations can significantly affect the reconstruction of cultural traits. The
thought experiment involves a group of individuals, one of whom is given a
picture that looks like a slightly crooked five pointed star; that individual is
asked to copy the picture and pass it, without showing the original, to another
member of the group, who is to copy it, and so on. Sperber claims that were
this experiment carried out, each individual in the group would draw a five

³ In essence, the point here is that factors of attraction may effectively form attractors that are
functionally and adaptively equivalent to the original cultural variation.
pointed star, and would not attempt to copy exactly what they had initially been given—if Sperber intended them to copy his drawing exactly, that subtle variation would be lost. While I think Sperber is right that this would happen, and is right that it would happen because of the learners’ expectations, the thought experiment departs from real teaching situations in some very important respects. In real learning situations, teachers often guide and correct their students, which gives the students more information about how to interpret the teacher; the range of things the teacher could mean given considerations of relevance gets repeatedly narrowed, and can be narrowed sufficiently to help them reconstruct even subtle or unexpected things. If the copiers in the group were told that they were to copy the drawing exactly as it appeared Sperber would get a different result than if they were simply given a star to copy. Hence, there are some reasons to think that learners do have significant assistance in acquiring precisely those features of cultural traits they need to acquire if cultural evolution is to be adaptive and cumulative.

3.2 Cultural selection may counteract the affects of attractors when cultural variation is adaptively important

In this section, I want to discuss one of the responses provided by the DITs themselves to the problem that attractors apparently pose for adaptive, cumulative cultural evolution (Henrich et al. [2008]). The claim is that factors of attraction greatly increase the probability that a cultural trait will be copied in the direction of an attractor, and over time cultural variation will drift to those points in the space. Suppose individuals are choosing who to copy from at random, and suppose there is some new innovation, trait T and an attractor A. Then if attractors are strong (i.e. the probability of a trait being transformed in their direction is high), then as individuals copy T at random, they are likely, on average, to end up with a cultural trait that is further from T and slightly closer to A. With each successive round of copying, the tendency in the population is to trend closer and closer to A. This is because in each generation each individual is just picking a teacher at random and hence is expected to end up with the average trait; and because each individual is then expected, on average, to transform that trait closer to A than it was.

The dual inheritance theorists generally argue that, if culture is to evolve by natural selection, it must be adaptive, and if it is to be reliably adaptive then there must be psychological mechanisms that bias humans in favor of copying more adaptive than average traits. Henrich and Gil-White ([2001]), for example, argue that humans have a tendency to copy cultural traits from prestigious individuals in their group, and this can lead, more often than not, to them acquiring traits that are more adaptive than average. The reason for this is that groups of humans tend to confer prestige on individuals where the
members of the group judge those individuals to have highly successful cultural traits (e.g. good hunters have more prestige than poor hunters) and the success in many of these cases correlates quite well with fitness. Suppose some group G tends to confer prestige on an individual A when A’s cultural traits are highly adaptive in some area. Then B can use that prestige conferred by G as an indicator of the fitness of A’s traits, and choose A as a model for her own cultural traits. Another possible heuristic for choosing fit cultural traits is a success-based heuristic—e.g. copy from the hunter that kills the most deer, etc. (Richerson and Boyd [2005]). Now if humans really do use such heuristics in choosing who to learn from—and the authors just cited provide some evidence that they do—then the problem posed by attractors is ameliorated. The reason is, given that T is more adaptive than A, that humans engaged in cultural learning will be using heuristics that make them tend, in each generation, to copy traits that are closer to T than to A, or (depending on how reliable these heuristics are) the trait that is closest to T. This means cultural selection will be pushing back against the force of the attractor—the better individuals are at choosing teachers close to T, the stronger the force acting against the attractor becomes. This is helped if multiple individuals can copy from the same teacher (as they can in human learning). This in itself, however, will not prevent the loss of T. However, there is a solution to this problem: attractors are probabilistic, so ‘error’ in cultural lineages in the presence of attractors, while tending towards the attractor, will always show some ‘scatter’. This means that occasionally an individual will generate variation that is further away from the attractor and closer to T; such individuals will then be available as models to the rest of the population and can allow that population to maintain variation closer to T than would otherwise be expected. Precisely, how good this is depends greatly on the ‘strength’ of the attractor, the degree of error, and the relative reliability of cultural selection. For Sperber’s own discussion of the interaction of cultural selection versus attraction in determining the outcomes of cultural evolution see (Claidière and Sperber [2007]).

Perhaps most importantly, cultural selection operates in favor of adaptive traits and, therefore, cultural selection is most likely to be an important force stabilizing and maintaining cultural variation in precisely those cases in which the dual inheritance theorists are most interested—the cases of adaptive cultural traits. Factors of attraction will probably be more explanatorily important in cases

---

4 I should emphasize again that Sperber and his colleagues’ own concern with the dual inheritance theorist’s claims about the power of cultural selection is how good a job such explanations do at explaining the ‘macro stability’ of culture, not as some have supposed, about the corruption of cultural variation. Sperber is not concerned if it turns out that cultural stability is partly explained by attractors and partly by cultural selection like forces; indeed, he seems open to this possibility (Claidière and Sperber [2007]).
where the trait is of relatively neutral adaptive value (and hence those cases in which Sperber and his colleagues are most interested).\(^5\)

### 3.3 Attractors may be numerous

So far I’ve argued that attraction does not necessarily mean that adaptive cultural variation will be lost in populations. This is because learning situations may be of the right type to permit individuals to correctly reconstruct at least the adaptively relevant features of cultural traits even in the presence of attractors, and because cultural selection is likely to be acting against attractors in the case of adaptive traits. However, even if individuals do not have psychological or teaching help to preserve the functional and adaptive features of the cultural traits they learn it may still be the case that they reconstruct cultural traits adaptively and functionally equivalent to the originals (i.e. having the same causal effect for the individual, or at least an adaptively indistinguishable causal effect for the individual, as the original trait) and hence preserve adaptive variation when it does arise, if there are a lot of potential attractors in the space of cultural traits. Cumulative cultural evolution requires that humans be able to transmit cultural variants and build on them, with each new variation improving on the previous variant. But even if there are attractors, and attractors are very strong, the more attractors there are the more likely it is that there will be attractors at points in the space sufficiently close to the initial cultural variation that movement toward that attractor leaves a trait that is adaptively and functionally equivalent to the original. As I have already discussed, this adaptive and functional equivalence is enough to permit adaptive cumulative cultural evolution of the type that the dual inheritance theorists want. What’s more, if there are a lot of attractors, then there are likely to be attractors close to cultural variants that represent accessible cumulative improvements on current adaptive cultural traits. So, even if cultural variation inevitably ends up at attractors, there can be cumulative cultural evolution that involves movement from one adaptive attractor to another that is more so.

So what reasons are there to think there would be a lot of attractors? There are a number of reasons why; some of these will be covered in the following sections. The first is where attractors are not universal (i.e. where there are significant differences between individuals in the factors of attraction acting on them). Human beings differ in the details of their psychological and social situation, their capacities for interpretation and understanding to the point

---

\(^5\) This also means that cultural selection is probably the most important force-generating stability—Sperber’s concern—in the case of adaptive traits, and attractors the most important force generating stability in the case of adaptively neutral (or maladaptive) traits.
that different individuals may have different factors of attraction acting on them and hence different attractors. If individuals have very different attractors, then population wide, at least some individuals will be likely to retain a functionally and adaptively equivalent version of the original variation (I’ll discuss this in much more detail in Section 3.5.) Another possibility is that there may be a wide range of factors of attraction acting on individuals, and more factors of attraction generally means more attractors towards which a learner might transform any new cultural variant. The more attractors each individual has for each trait, the more likely it is that the closest attractor to any new cultural variation is functionally and adaptively equivalent to it; and hence that each individual will end up with a functionally and adaptively equivalent variant after reconstruction. Another reason to expect many attractors at the level of cultural traits may be if factors of attraction are essentially acting at the level of the components of cultural traits; I will discuss the reasons for thinking this might be the case in the next section.

3.4 Factors of attraction may often be working on the components of cultural traits

In this section, I want to discuss another way in which cultural evolution might still manage to be adaptive and cumulative even with very strong attractors: this might happen where factors of attraction, and hence attractors, determine the reconstruction of the components of cultural traits rather than the cultural traits themselves. What do I mean by this? Many cultural traits, especially complex and adaptively important behaviors, artifacts or ideas are readily decomposed into component movements, actions, pieces or ideas; and, indeed, teachers often deliberately emphasize such components when teaching as a means of assisting learning. For example, consider a case of a possibly adaptively significant cultural trait: the throwing of a pot. Throwing a pot involves a series of steps, such as beating the clay to remove air bubbles, transferring clay onto a wheel, operating the wheel, the various steps involved in shaping the pot on the wheel and the hand motions necessary to do it, etc. In order to acquire the cultural trait, the learner needs to acquire each of these steps correctly and assemble them in the right order; in order to get variations in the cultural trait there will have to be substitutions, additions, or subtractions of such steps. The same goes for some of the less obviously adaptively significant traits to which Sperber refers: for example, little Red Riding Hood or Dawkins’ origami junk; these traits have steps or components (e.g. folds, ‘chunks’ of the story).

Notice, furthermore, that in some of the basic examples to which Sperber refers, it is not the overall cultural trait which is subject to attractors, but the components of those cultural traits. In the case of Dawkins’ origami junk, for
example, the thing that is corrected by the learner boys—and which is thereby subject to an attractor—is not the whole ‘origami junk’ trait, but a component fold (e.g. ‘fold the paper into the exact middle’ rather than ‘fold the paper crooked by 1.5 degrees and a millimeter from the middle’). If this is true of other cultural traits, then in some cases factors of attraction will be acting to determine how the learner reconstructs individual components of a trait, rather than exactly what the overall trait looks like. One obvious reason why this might happen is that teachers often emphasize the modular character of cultural traits. By doing this, they may make it relatively easy for the learner to appreciate that there are a series of steps in the correct overall structure of the cultural trait; however, this shifts the learning problem from how to correctly reconstruct the overall structure of the cultural trait to how to correctly reconstruct each of its components. In such cases, factors of attraction would then effect how the learner reconstructs these components.

However, if it is often the case that attractors are acting at the level of components of cultural traits, then, even with strong attractors, cultural variation could be very rich because cultural variation could be ‘built out of’ these attractors. Human beings may prefer precise folds over slightly off kilter ones, but this still permits a vast array of possible origami items, including cumulatively elaborated ones: even if humans are more likely to remember slightly formulaic elements in stories, there can still be a vast array of stories, even cumulatively elaborated stories. In fact, many artistic traditions (especially folk art traditions) work precisely this way—they are built out of elements that are easy to remember, elements that are essentially attractors. This does not prevent these traditions from becoming very rich. To make this point clearer, let’s return to the case of the pot. Suppose there are ‘attractors’ for all the various steps that are involved in making a pot—e.g. hand motions easier to remember or perform—then these components will be drawn to these attractors as the technique is transmitted. Nevertheless, so long as the components’ moving to attractors doesn’t significantly change the functional properties of the final product, this will make little difference to the possibility of cumulative improvements to this trait. Cumulative cultural change will then largely have to involve adding, removing, or substituting steps rather than making subtle changes to the steps themselves. But in most cases this would still allow a significant cumulative cultural change to the item in question—for example, pots can be elaborated in various useful ways—changed in shape or

Of course, attractors at the level of components will still effectively generate attractors at the level of whole traits, but these will be very numerous (i.e. there will be a trait-level attractor at every point in the space equivalent to a collection of component-level attractors)—and this itself will significantly lessen the problem posed by attractors (as described in the previous section).
dimension, given handles or spouts, etc—simply by adding or substituting steps in the process, changing tools, and so on.\(^7\)

What’s more, if reconstruction is something that happens to the components of cultural traits, then one can get significant cumulative cultural evolution even in cases of maximal constraint on the reconstruction of those components. Sperber ([2000]) imagines reconstruction as lying somewhere between two extremes of perfect copying on the one hand (where all of the information in the cultural trait comes from the copied trait) and triggering on the other (where all of the information in the ‘cultural trait’ comes from what is already present in the learner’s mind). Sperber ([2000]) offers two examples to illustrate the difference between copying and triggering—the music players. Imagine a line of music players, and that these music players are also music recorders. This means that when the first one plays a piece of music, the second player records the music of the first one, and then plays that music back; the third one then records the music of the second player, and then plays it back, and so on. In this case, according to Sperber, the players are really ‘copying’ the music. Now imagine another line of players. This time, the first player plays a tune stored in its memory; this leads the second player to play the same tune, which is also stored in its memory; which leads the third player to play the same tune from memory, and so on. In this case, according to Sperber, it appears the music players are simply triggering representations. However, if what are reconstructed are the components of cultural traits, and not cultural traits themselves, then even if reconstruction were essentially triggering, cumulative cultural evolution would still be possible. To see why this is so, consider a third case: imagine this time that there is a line of music players that are restricted to reproducing a few dozen notes. When the first player plays a piece of music, the second player plays a sequence of notes as closely corresponding to the notes of the first player as possible (it will be restricted by the range of notes available to it); the third player then plays the same sequence using its available notes, and so on. Add in some mechanism for variation and selection, and this case does seem to permit cumulative cultural evolution of a sort, with the machines transmitting simple tunes composed of those notes; and the more notes the machines can play, the greater the potential for ‘cultural’ variation and ‘cumulative cultural processes’ that they have. Despite these music players being maximally constrained to reconstruct the music they hear in terms of the notes available to them, there is much less constraint at the level of the entire tunes.

\(^7\) Note that this view of how cultural variation works allows for significant but nevertheless constrained cultural variation and thereby explains both stability and cumulative cultural evolution, a major desideratum in this debate.
3.5 Attractors are highly variable, even within local populations

The central problem that attractors are supposed to pose for cumulative cultural evolution is that they generate *directional* transformation in cultural traits, rather than simple error—as I discussed in Section 3.2, simple error can actually help to permit cumulative cultural evolution in the presence of cultural selection. Sperber’s simplest examples of attractors involve the transmission of a trait through a lineage, with each individual (or most individuals) in that lineage having a tendency to transform that trait closer to an attractor—the same attractor—and hence the cultural trait is eventually pulled to it. In these cases directional transformation occurs because the individuals in the lineage share psychological, social, and ecological factors of attraction that lead them to transform the trait (generally) in the same direction. So much of the *prima facie* problem posed by attractors stems from attractors’ being universal—if individuals in a population all tend to share the same factors of attraction, then loss of cultural variation is inevitable. In this section, I am going to argue (as, indeed, I think Sperber clearly believes (Sperber [1996], pp.115–17)) that attractors are not *always* universal; more importantly I will explain why their not being universal means that cumulative cultural evolution can still occur. I’ll consider two main types of ways in which cultural variation can persist in the face of non-universal attractors.

The first type of case where attractors for some cultural trait would not be universal is where attractors differ by sub-population: i.e. where at least some individuals either do not have or have weak attractors with regard to some types of cultural traits compared with the rest of the population. (This might happen due to differences in the capacities, knowledge, preferences, training, etc., of the individuals in question from others in the population). If this happens then cultural traits will be retained and transmitted by these individuals in close to their original form. This will slow, but not prevent the loss of that cultural variation. However, if such lack of (or weaker) attractors is not *very* rare, and occurs at least occasionally in each cultural generation such that similar individuals are at least occasionally transmitting these relatively untransformed traits to each other, then, in effect there is a *sub*-population of individuals capable of retaining the cultural traits in question even if they are lost or drawn to attractors in the rest of the population. Take, for example, professional wrestling moves. Now, wrestling traits transmitted in the general population fail to persist because most individuals don’t have the interest or capacity to retain and perform them, and hence these traits would probably rapidly disappear or be drawn to attractors. Nevertheless, because professional wrestlers have the training and capacity (and preferences) that permit them to retain those moves relatively unchanged, and because they transmit those moves to each other, professional wrestling moves are maintained in the
professional wrestler sub-population even if they are lost from the rest of the population. In other words, if humans vary with regard to the factors of attraction acting on them such that some individuals don’t have or only weakly have attractors with regard to at least some cultural traits, then attractors cannot remove those traits from the population as a whole. What’s more, there is some empirical evidence that such sub-populations are important for the retention of different sorts of complex cultural traits within real human populations. There are cases in cultural history where this has been a problem—for example, the massive reduction in the hunter gatherer population of early Tasmania led to a loss of a significant number of complex cultural traits (Henrich [2004]), probably because smaller populations retain less variation—and hence have fewer individuals with the skills necessary to retain and pass on any particular type of trait. These small numbers of individuals are much more likely to be lost by accident, and take with them their capacity to retain the traits in question.

Sperber and his colleagues clearly accept that attractors are certainly variable across historical time (Claidiere and Sperber [2007]) and somewhat between individuals (Sperber [1996]); hence, their view is compatible with the above type of case and hence allows for adaptive, cumulative cultural evolution. However, in practice, factors of attraction may often be much more variable even that this—they not only vary by sub-population, but may vary greatly by individual and even situation and operate in different combinations in different cases. This means that attractors will vary with those different contexts, too. To give some examples of such situation variable constraints, let’s go back to the case of throwing the pot or folding the origami junk. One obvious constraint on how the origami junk is copied is the background knowledge of the boys who are learning the trait. For example, in the origami case, folding exactly in half is only an attractor because the boys already know something about origami and that ‘folding exactly in half’ is a standard step in origami whilst ‘folding off center by one degree’ is not. Background knowledge about origami is obviously not a feature that all humans share—it’s not even necessarily a feature that all the boys themselves share. Consequently, an entirely different sort of cultural trait involving paper folding might not lead to the same attractors if the boys’ knowledge about how that sort of trait was to be performed was different.

Another factor of attraction that is highly variable between individuals is the learner’s degree of knowledge about the functional and causal properties of the items or behaviors being taught—for example, let’s go back to the potter. Imagine the teacher is demonstrating the hand movements that

---

8 I should note here that Henrich is not assuming variation in skill due to variation in learning capacity, but simply random variation due to learning error. However, in principle, psychological variation affecting how well a trait was picked up should have much the same effect.
shape the pot on the wheel, and as she is doing this her hand twitches slightly, causing a furrow in the side of the pot. When reconstructing the teacher’s behavior to make their own pots, the learners are likely to leave out the twitch, because they can see that the furrow actually disrupts the functional properties of the pot. Again, the learners probably have gained this appreciation of how pots are supposed to work from their past experience with these pots and from their previous learning process with their teacher; an appreciation of how these pots work varies greatly with individuals, and what’s more, how well they apply that appreciation in different contexts; not all pottery learners generally, nor even in the same population, will apply that knowledge in the same way. Nevertheless, it determines how the learners in this case reconstruct the teacher’s pot making. In other cases, one could imagine different learners coming to different conclusions about whether to leave things in or out of their reconstruction based on what they believe are the relevant functional considerations.

In the case of the transmission of ideas, another parallel candidate is the relative reasonableness, logical consistency or common sense nature of the idea, and the ease with which such ideas can be synthesized with the rest of the individual learner’s knowledge. Sperber and his colleagues in the EoR group have noted that humans tend to easily remember counterintuitive ideas—probably because they violate expectations—but they have also noted that the presence of such violation shows that in normal cultural transmission individuals expect such rules to apply to ideas that are transmitted, and will most naturally interpret what is communicated to them as ideas fitting into their current understanding of the world. Indeed, counterintuitive ideas are only easier to remember only where they are part of a larger idea that is otherwise intuitive—such as ghosts being invisible or able to interpenetrate objects but still having normal human motivations (Boyer [2001]). Which ideas will actually be ‘attractors’ here will depend on characteristics individuals share—such as the human ‘species typical’ psychology—but also on the beliefs considered normal in the surrounding population and, indeed, on the individuals themselves (more skeptical individuals may be less likely to accept claims that the noises in the forest come from ghosts, for example; individuals will also vary in what background information beliefs have to fit in with).

Finally, and perhaps most obviously, there is the constraint of what the learner thinks the teacher herself intends to be part of her action. With the hand twitch, above, the learner potter might conclude that the teacher didn’t intend this by seeing that the twitch looked spasmodic, the sudden frown on the teacher’s face, or the expression of irritation or disgust. In other cases the clues might be more subtle. Some of this information might come from ‘species typical’ psychology (for example, domain-specific knowledge of how human minds work), but some of it is determined by the learner’s
understanding of the teacher’s personality and the learner’s capacity for accurate interpretation of others. What’s more, as discussed in Section 3.1, a teacher often spends a significant amount of time with a learner, offering feedback. This can often mean that a teacher can correct a learner’s tendency to interpret them in one way or another. This in itself could affect the tendency of some ‘attractor’ to appear.

The existence of all these different types of elements constraining the reconstruction of the cultural trait—and working together in different ways in different combinations in different cases—mean that what actually ends up being the ‘attractor’ may be different in almost every individual situation (and thus not really be an attractor at all in Sperber’s sense of a force acting on an entire lineage or population). In cases where such individually variable factors of attraction are very important there will be insufficient consistency in forces of attraction to cause population level ‘pull’ towards one kind of trait; and this in turn means that transformation that occurs in a cultural trait is random—in the sense that it is controlled by factors largely external to the cultural transmission process. And as previously discussed, a certain amount of random transformation is not a problem for cumulative cultural evolution so long as there are adaptive learning heuristics such as prestige or success biases being employed by the learners: this is because random learning error generates additional, potentially adaptive cultural variation that learners with such biases can selectively acquire. So if forces of attraction are variable, as Sperber believes they often are, then attractors will not consistently destroy cultural variation; where this varies by sub-population, then cultural variation can persist and become cumulatively adaptive within those sub-populations.

In cases where the factors of attraction described above are very individually variable, then they will generate no consistent force in lineages to pull a cultural trait in one direction in a population—transformations of the trait will be essentially random or unpredictable. So variation in attractors prevents attractors from working against cumulative cultural evolution.

4 Conclusion

The dual inheritance theorists have worried that Sperber’s theory of attractors, which was primarily designed to explain the stability of cultural traits in the face of errors in transmission, must also show that cumulative cultural evolution is impossible. I think I have offered a number of reasons not to fear the consequences of Sperber’s view for cumulative cultural evolution: in most cases this is because the sorts of attractors he is proposing are not of a kind to prevent such processes from occurring. There are a variety of misunderstandings between the EoR group and the DITs that might explain why attractors look problematic for adaptive cumulative cultural evolution, for example.
misunderstandings about what each group has in mind by the claim that cultural transmission is or is not error prone and the different emphasis and paradigm examples on which the EoRs and DITs focus (the dual inheritance theorists are concerned about determining how cultural evolution can be adaptive in order to explain how cultural psychology might have evolved; Sperber and his colleagues are more interested in explaining stability in cultural traits, often emphasizing neutral cultural traits). However, these misunderstandings are just that: ultimately, I think the EoR’s and DIT’s views are compatible and they are converging on a common story about the nature and evolution of culture.

Acknowledgments

Many thanks to John W. Carroll, Dan Sperber and several anonymous reviewers for their very helpful comments. This paper is much better for their input.

References

Catherine Driscoll


