

An agnostic view of memes

Robert Aunger

LSHTM

It has long been noted that there is a tendency for the beliefs and values of one generation to resemble those of the previous generation. It is widely thought that this is due to a kind of inheritance process, whereby the values of a group are passed from one time to the next as a way of preserving their “culture.” There are (at least) three causal mechanisms that can explain the recurrence of cultural traits (Aunger, in press). Recurrence can occur through 1) strong individual learning biases; 2) population-level normalizing effects on the traits which are adopted; and 3) replicator-based inheritance. Evolutionary psychologists (e.g., Tooby and Cosmides, 1992) tend to believe that the first mode is most likely by suggesting that much of what appears to be socially learned (i.e., ‘culture’) is in fact ‘evoked’ from an individual’s genetic stores of knowledge by environmental stimuli. Cultural evolutionists (e.g., Henrich and Boyd, 1998) emphasize the second possibility, that conformist transmission, or the tendency to do what others are doing, is responsible for inter-generational cultural similarity. The clear position that the field of memetics takes – the point that separates memetics from all the other positions on cultural evolution – is that social learning primarily involves the replication of particles of information, called ‘memes’ (Dawkins, 1976; Dennett, 1997; Blackmore, 1999; cf. Gil-White 2004). Thus, memetics is associated with the third possible means of accounting for cultural continuity. The central hypothesis of memetics is the claim that cultural replicators (memes) are duplicated through the process of social learning, particularly imitation. My strategy in this paper will simply be to work out what this implies, to take the meme-as-cultural-replicator position as far as it will go – but no further. I don’t assume (unlike the dominant memeticists, Dennett 1997; Blackmore 1999; Distin, 2004) that memes exist. In fact, “meme,” as it is currently used, is a concept I consider too diffuse to be subject to scientific testing, much less to support the claim that memes have played a crucial role in the evolution of large brains and human society. (Blackmore 1999)

The several schools of thought noted above all agree that cultural change is an evolutionary process (see Laland and Brown 2002 for an overview). What they disagree about is the mechanism. What most people don’t realize – including distinguished figures in evolutionary theory, such as Richard Dawkins (1976) – is that evolution doesn’t depend on replication. A Darwinian process is defined by descent with modification (Lewontin 1970). It so happens that the quintessential, prototypical evolutionary process, the origin and regeneration of species, is based on the inheritance of genes, which are replicators. But for cultural change to be an evolutionary process, it isn’t necessary to hypothesize that memes exist. Cultural evolution could be based on a blending inheritance rule, such as averaging the values exhibited by multiple cultural ‘parents’ (Boyd and Richerson, 1985). Blending inheritance has been proposed as a model of biological evolution as well: Galton (1898) proposed (before Mendel’s law was rediscovered) that an individual’s genetic makeup was a combination of contributions from parents, grandparents and other ancestors. This process of convergence on the mean would still be Darwinian, because it exhibits descent with modification, but it would not preserve the identity of particular bits of information through transmission.

Most observers agree that cultural information is inherited through a process of social learning. Social learning in humans often involves communication, verbal or non-verbal. But our best current models of communication (Sperber and Wilson, 1995; Grice 1989) are consistent with any of the three models of cultural change outlined above. We therefore can't just study the way people talk to each other to discriminate between the evolutionary psychological, cultural selectionist and memetic approaches to cultural evolution because communication models are not specific about how the brain treats information it acquires from social exchange. We have to know more about how communication works – how what I have in my head comes to resemble what you have in your head – before we can argue that the process is one in which a particular bit of information gets replicated, and hence can confirm that memes exist.

To determine if memes exist, it is also necessary to have a firm idea of what replication involves. Replication can be defined as a special relationship between a source and a copy such that four conditions hold:

- causation (the source must play some role in bringing about the conditions that lead to a copy being made);
- similarity (the source and copy must resemble each other in relevant respects);
- information transfer (what makes the copy similar to the source must be derived from the source); and
- duplication (the source and copy must coexist for some time). (Aunger, 2002)

In the Dawkins/Blackmore/Dennett tradition, cultural information can be present in any behaviour or artifact, as well as in minds, because people can extract that information from any of these forms and copy it for themselves (e.g., through a generalized notion of imitation, *sensu* Blackmore 1999). In standard memetics, then, replication is considered to be a much more general process than defined above. This allows memes to be instantiated in behaviour, artifacts or ideas.

Other memeticists argue that memes are restricted to particular kinds of physical forms. For Distin (2004), for example, memes are restricted to representational objects. Thus, some artifacts, such as books and CDs, can house memes, but non-representational artifacts such as everyday objects (spoons, cars) cannot. Some behaviour can be symbolic, as when it is ritualized (e.g., dance), or when two lights flashing on a hilltop mean the British are invading (as in the Paul Revere myth), and thus embody memes, while everyday behaviour cannot. In this case, memes can exist in many forms, but they are all symbolic forms, so this is a more restricted manifestation than standard memetics, because most behaviour is not symbolic, and only a few kinds of artifacts are. The same basic problem remains, however: that memes must be able to replicate from any given symbolic form to any other. Thus, memes can move from a mind to a CD and back again.

It is true that any replicator can be thought of as information. Of course, information must take a form, or have a material instantiation. But in many cases, there can be multiple representations of a given bit of information. For example, it is possible for a gene sequencer to store the information about a particular gene in a computer, which is then reconstituted using different source of DNA by another machine. So is it legitimate to think of a gene being replicated in this fashion? What about the following example. Say a piece of paper on which a phrase is typed is scanned, then the phrase is converted to a character sequence by character recognition software, and then stored in computer memory, later to be spoken aloud by another piece of software, so that someone can hear the phrase and store it in their brain. Do the piece of paper

and the brain have the 'same' information? And has the 'same' meme gone from a piece of paper to a brain – is this an instance of replication?

These are central questions for memetics. Dawkins, Blackmore, Dennett and Distin answer 'yes' to both questions. For all of these scholars, the crucial fact is that symbolic information has a Protean ability to take multiple forms without its essential nature being changed. The fancy software, lab equipment and human activity involved in producing the gene copy or meme copy in these cases can be thought of as very complex replication-assistance machines. So a meme with content 'hello' can be in the representational form of words on paper, a spoken phrase, or the idea in someone's brain. Identifying a *particular* meme (a token of a given type) only requires being specific about its informational content and the form in which it is represented.

I argue that there are dangers in taking this 'pure information' route. Why isn't it possible to argue that memes exist only in brains and that these other forms are signals of various kinds (even written text, or digital memory)? Even if the information can be considered to be the same, *only in certain contexts does the same information have the ability to instigate the process through which its content is duplicated*. For example, only if the physical chain of amino acids is present inside a cell, with everything that implies, can a DNA molecule reproduce itself. The same is true of other replicators: the protein replicators, prions, require contact with other prions inside bodies; computer viruses only work when present in computers using a specific operating system. The process by which a meme replicates itself can become very complex indeed, involving transfers to artifacts of various kinds (as in the example above), for varying lengths of time, prior to finding its proper form again in another brain. A pattern of ink on paper can be the source material, the template, for multiple replications. However, it cannot cause itself to be used as a template. Genes have this unique ability only when inside organisms: they can produce proteins which then read off the gene sequence and construct a copy of that gene.

In essence, my argument is that a replicator can take many forms, but only one form is active; the other forms are inert. Thus, given the criteria for 'true' replication above, which includes an ability to play a causal role in the duplication of information, it turns out that replication is a sufficiently complex operation that it cannot proceed on multiple substrates, in multifarious contexts, as supposed by most memeticists. As a result, memes cannot be produced in the form of behavior, ideas, and in artifacts like books or TV shows (contra Dawkins, 1976; Dennett, 1995; Blackmore, 1999).

If memes could exist in brains, in speech and in artifacts, they would be the superheros of the replicator world, able to transform themselves into any shape or form at will, like the Proteus of Greek mythology. Instead, memes must be confined to one physical substrate, just as their brethren, the biological replicators genes and prions, are. I thus argue that only one substrate can be associated with memes.

So which substrate is the proper 'home' of memes? Artifacts do not satisfy the conditions of replication because they cannot transfer information directly from one to another; behaviour fails the replication test because it doesn't duplicate (an imitated behaviour usually occurs after the model behaviour has been displayed, so there is a sequence of one behaviour, then another). (Admittedly, there are exceptions: computer viruses are artifact-based replicators, and some behavioural mimicry leads to the same behaviour being simultaneously exhibited by

two people, but these are rather exceptional exceptions to the rule.) The only remaining candidate is that memes are mental objects.

If memes can only take one form, and its physical substrate is tied to brains, then how can it be that memes make their way from inside one brain to another? Brains are not in direct physical contact with one another, so memes have to take some intermediary form to ‘jump the gap’ between people -- something like a spoken phrase. I propose that meme replication solves the ‘gap problem’ through a process I call ‘instigation’. Memes, as brain states, produce motor behaviour that produces signals, such as spoken language, which, on being heard, initiate a complex process in the receiving individual’s brain which culminates in a new brain state that qualifies as being the same as the memetic state which produced the signal in the first place. The role of signals in this view is not to bring along the information that will be incorporated into the meme copy, but rather to set in motion the kind of local transformation that will result in the copy being produced. The signals themselves are not memes, nor need they contain representations of memes (i.e., the same information as the meme that produced it). Instead, the signal contains the information necessary for the signal-receiver’s brain to reconstruct the source meme *in situ*. This may be some ‘trigger’ which relies on innate (and hence shared) structures in the receiving brain which can be depended on to do the job of reproducing the relevant memetic information from that trigger. Thus the content of the signal does not necessarily map in any way onto the content of the meme. I thus define meme replication as a process in which such neuronal states are duplicated through a process instigated by signals (Aunger, 2002).

Alternatively, brains can instigate a process that leads to the production of a sign -- typically a sequence of behaviours that result in an artifact which serves as an intermediate store of information related to meme production. When observed by a naïve individual, this sign produces signals that can lead the observer/receiver’s brain to go through the reconstructive process just described. Books and other representational artifacts thus play a crucial role in the evolutionary dynamics of modern culture.

This perspective has a number of advantages. First, it adheres to my argument that memes can only be in one place: people’s heads, not anywhere else. Other memeticists argue that information is still a meme when on paper or computer screen, even though it isn’t capable of initiating action independently. I argue, on the contrary, that the pattern of writing on paper is a ‘signal template’ – a stable substrate from which signals can be created (Aunger, 2002). For example, photons can glance off a piece of paper, becoming ‘informed’ about the content of the writing on it in a way that is interpretable by human brains as a signal; this sensory signal in turn, once inside the brain in the form of inter-neuronal signals, instigates meme replication by transforming the state of some neurons into a state similar to that of an area in the signal sender’s brain.

Second, it side-steps Sperber’s (2000) complaint that communication is too imprecise a process to result in replication, and is consistent with the ‘poverty of the stimulus’ postulate in communication theory (that the amount of information contained in messages is insufficient to define the inferences made on the basis of those messages). Signals are not memes, so the content of communication can be different from the consequences of it. As a result (third), it keeps memetic evolution Darwinian because what happens to signals (Lamarckian influences on memetic phenotypes) has no necessary influence on the ‘germ-line’ of meme copies in brains. The information content of offspring memes isn’t affected by anything that happens to degrade or modify signals because they are built out of local materials in the receiving person’s brain. The signal’s message can just be “Go to work to produce meme X!”; nothing more.

Fourth, it keeps the account of memetic ‘generations’ clean: from brain to brain constitutes one generation, no matter how complex the intermediary process. Fifth, it keeps the evolutionary lineages clear: replications of a meme can be traced by ‘counting heads’, not artifacts, or behaviour or anything else. So memes can persist, form long lineages, and explain cultural similarity over time. In effect, *all* the hurdles I believe meme theory must jump to explain cultural transmission are leapt with this one conceptual move: seeing signals as instigators of meme reconstruction rather than as alternative embodiments of memes themselves.

But what is the likelihood that information replicates in brains? The brain dynamicist school (e.g., Freeman 1999, McCrone 1999) believe that it’s basically senseless to break the brain into bits. In their lingo, borrowed from complexity theory, the whole cortex flexes when stimulated by new data, then eventually settles down into a state of relaxation that represents the individual’s refined response to some situation. The idea of “active” memes is antithetical to such a view because the meme argument requires that some parts of the brain be able to function as individualized, independent units of information storage, transmission and replication.

For brain dynamicists, it is also difficult to see why information should be duplicated in the brain, because they think of the brain as being involved in the production of a single, complex representation of everything you know or feel or think, and that huge complex changes every second. But if information is in any sense localized in the brain – if you can say that particular areas play important roles in the retrieval and representation of, say nouns or the ability to recognize faces (as brain scans nowadays seem to suggest) – then the role of particular nodes in the storage of that information for future use must be guaranteed. But a single node is degenerate source – neurons die all the time. Backups become necessary, and that means more than one copy of the “same” information, which presumably occurs through a process of replication (just what memes are supposed to do), would have utility. Other eminent brain scientists are quite happy to think of the brain as composed of individuated circuits (Fuster 2002 or Mountcastle 1998), more in concert with the position of a memeticist.

The replication model I have suggested for memes doesn’t prove that memes exist; it just gives them a way to exist. The model may make sense, but that doesn’t force the objects of the model to be real, to be things-in-the-world. Given the current state of ignorance concerning how brains store and manipulate information, an agnostic position on the existence of memes is the only scientifically responsible position to take. My objective has simply been to make a diligent attempt to take what is a very poorly understood concept more scientific, and in particular, susceptible to empirical testing. I’m content to challenge people more competent than myself to settle the matter one way or another.

I think it is quite possible that the brain doesn’t work in the way it must do to be a store-house and replicator of memes. In this case, then I think we have to admit that cultural evolution proceeds by another mechanism than replication. If the model I have sketched here – or a similar one consistent with the constraints I have developed for any model of meme replication – is not empirically validated, then the idea that memes exist must be abandoned because alternative approaches to meme replication have already been eliminated. That is, we will no longer have any reason to refer to memes in our explanations of socio-cultural phenomena. I suggest this is a contribution to science since it would eliminate one hypothesis for how culture works. We will then be left with the interesting prospect of having to develop a new evolutionary theory of cultural change that does not depend on the replication of bits of information. Culture then would not be like biology in important respects. This conclusion, it seems to me, may present its own problems since gene-culture coevolutionary models would then be attempting

to simultaneously describe one process that conforms to the replicator dynamic (genetic evolution), while the other does not (although Henrich and Boyd 2002 have shown that a strong attractor-style transmission process works very much like a replicator process in terms of population dynamics).

Is this theory rich enough to explain the complexities of human culture? Probably not. Memes are just parasites on the normal system of thinking and communicating. Memes can't explain big topics like the evolution of big brains or human culture (contrary to Blackmore 1999), because not all thinking replicates information, nor does all communication. So memes are the result of a potentially important (from an evolutionary point of view), but also a potentially empty, subset of cultural processes. Memetics, if it explains anything at all, can only explain how a parasite arose on the system for thinking and exchanging information in social species with relatively big brains. It is not, and cannot be, a general theory of culture.

In the end, then, it seems that worrying about the mechanism by which culture evolves may not be central to the explanation of how or whether culture differs from other evolutionary processes. It is also not central to the question that concerns (at least implicitly) most memeticists: how *human* culture differs from the culture of other animal species. So the issue of whether culture is underwritten by memes may be peripheral. Certainly, human culture has a very different evolutionary trajectory than non-human culture. No doubt any explanation of this difference must include reference to the greater ability of humans to communicate specialized information (i.e., language) and to learn socially (i.e., imitation). Either of these might be sufficient to explain why human culture is able to accumulate more rapidly than cultural traits in other animals. It also seems likely that the ability to store information in the environment, in the form of artifacts, is also crucial (Donald, 1991; Clark, 1998; Aunger, 2002). There's a lot more happening in culture than just the replication of information. Explaining human culture requires attention to the independent dynamics of technology, not just changes in beliefs and values, because technology increasingly has the power to transform the physical environment in ways which influence both human culture and biology (Aunger, 2002).

What roles do artifacts play in cultural evolution? Certainly, as described above, artifacts can serve as way-stations for information content related to memes, but do not serve as storehouses nor as creators of memes themselves, which are restricted to their natural place inside brains. Of course, artefacts have roles as stores of information (e.g., books), and as channels of information (the World Wide Web), whether or not the information involved is memetic. So artefacts are guaranteed members of the cultural evolutionary team while memes are only possible players.

Still, if memes exist, they have to be the cause of *some* important phenomenon – just like every other replicator operating in the world. Prions cause mad cow disease, computer viruses wreak havoc with the Internet; genes make us. If we root out memes from their hiding places, we will begin to understand how cultural change works, why people cling to certain beliefs that seem not to do them any good (because they're infected by brain parasites with their own interests), and have yet another example of a replicator to compare with those we already know about. This makes the meme project worth working on. I can't establish that anything associated with this system are memes, but it is still fascinating to think that someday, we might be able to tell whether real people in real situations are in fact spreading memes around when they interact. Memes will only be proven to exist when (if?) the lab-based (rather than theoretical) neuroscientists become interested in this question.

REFERENCES

- Aunger, Robert (in press) "Three roads to cultural recurrence," in *The Epidemiology of Ideas*, ed. by Dan Sperber. Chicago: Open Court
- Aunger, Robert (2002) *The Electric Meme: A New Theory of How We Think and Communicate*. New York: The Free Press.
- Blackmore, Susan (1999) *The Meme Machine*. Oxford: Oxford University Press.
- Boyd, Robert, and Peter J. Richerson (1985) *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Boyd, Robert, and Peter J. Richerson (2000) "Memes: Universal acid or a better mouse trap?" In *Darwinizing Culture: The Status of Memetics as a Science*, Robert Aunger (ed.), pp.143-162. Oxford: Oxford University Press.
- Clark, Andy (1998) *Being There: Putting Brain, Body, and World Together Again*. Cambridge: MIT Press.
- Dawkins, Richard (1976) *The Selfish Gene*. Oxford: Oxford University Press.
- Distin, Kate (2004) *The Selfish Meme: A Critical Reassessment*. Cambridge: Cambridge University Press.
- Donald, Merlin (1991) *Origins of the Modern Mind: Three Stages in the Evolution of Cognition and Culture*. Cambridge: Harvard University Press.
- Freeman, Walter J. (1999) *How Brains Make Up Their Minds*. London: Weidenfeld and Nicolson.
- Fuster, Joaquin M. (2002) *Cortex and Mind: Unifying Cognition*. Oxford: Oxford University Press.
- Galton, Francis. (1898) 'A diagram of heredity'. *Nature* 57:293.
- Gil-White, Francisco J. (2004). 'Common misunderstandings of memes (and genes): The promise and the limits of the genetic analogy to cultural transmission processes'. In Susan Hurley and Nick Chater (Eds.), *Perspectives on Imitation: From Mirror Neurons to Memes*. Cambridge, MA: MIT Press.
- Grice, H. Paul. (1989) *Studies in the Way of Words*. Cambridge MA: Harvard University Press.
- Henrich, Joseph and Robert Boyd (2002) 'On modelling cognition and culture: Why replicators are not necessary for cultural evolution.' *Journal of Cognition and Culture* 2(2): 87-112.
- Henrich, Joseph and Robert Boyd (1998). The evolution of conformist transmission and between-group differences. *Evolution and Human Behavior*, 19: 215-242.
- Henrich, Joseph, Robert Boyd and Peter Richerson (in press) 'Five common mistakes in cultural evolution'. In *The Epidemiology of Ideas*, ed. by Dan Sperber. Chicago: Open Court.
- Laland, Kevin, and Gillian Brown (2002/3) *Sense and Nonsense*. Oxford: Oxford University Press.
- Lewontin, Richard C. (1970) "The units of selection," *Annual Review of Ecology and Systematics* 1:1-18.
- McCrone, John. (1999) *Going Inside: A Tour Round a Single Moment of Consciousness*. London: Faber and Faber.
- Mountcastle, Vernon (1998) *Perceptual Neuroscience: The Cerebral Cortex*. Cambridge, MA: Harvard University Press.
- Pinker, Steven (1997) *How the Mind Works*. London: Penguin.
- Richerson, Peter, and Robert Boyd (2004) *Not By Genes Alone*. Chicago: University of Chicago Press.

- Sperber, Dan (2000) "An objection to the memetic approach to culture," in *Darwinizing Culture: The Status of Memetics as a Science*, Robert Aunger (ed.), pp. 163-74, Oxford: Oxford University Press.
- Sperber, Dan (1996) *Explaining Culture: a Naturalistic Approach*. Cambridge, MA: Blackwell.
- Sperber, Dan and Deirdre Wilson. (1995; 2nd ed.) *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Tooby, John and Leda Cosmides (1992) "The psychological foundations of culture," in *The Adapted Mind*, Jerome H. Barkow, Leda Cosmides and John Tooby (eds.), pp. 19-136, Oxford: Oxford University Press.