for resource commitment strategies – may be even more plentiful and powerful. And contrary to M&D's tempting suggestion, such positive illusions are not restricted to subjective beliefs that are "not likely to be rudely contradicted by experience" (sect. 14, para. 3). Many of these widespread beliefs entail almost delusional denials of repeated experience. Notions that Eros lasts forever, this time it's real, and (as the sappy song says) "When we're hungry, love will keep us alive" are effective and virtually ubiquitous catalysts for reproductive pairbonding. But by non-reproductive periods of the human life cycle, those Romeos whose romantic illusions have not killed them, have oft' yielded to the wisdom of Friar Lawrence: "These violent delights have violent ends, and in their triumph die. . .Therefore love moderatedly: long love doth so. Too swift arrives as tardy as too slow." Yet another category altogether, unexamined by M&D, is selection for cognitive extravagance independent of problem-solving utility (Miller 2000; 2001). But even granting M&D's conclusion that there are just a few families of adaptive misbelief, we don't yet know enough about their natural history to determine how many species there are or what their carrying capacities and competitive coefficients are relative to true beliefs.

Second, even if reproductively beneficent misbeliefs are rare and most misbeliefs have costs, this does nothing to tell us how well evolution ultimately avoids such costs. Indeed, M&D elegantly acknowledge that functional normativity does not entail statistical normality: In evolution, forgivable malfunctions may be common and achieving proper function may be "positively rare" (sect. 3, para. 5). Thus, even if truth is the evolutionary target as M&D maintain, design constraints, by-product associations, and historical contingencies may make it one that cognition has a low probability of hitting.

Third, many kinds of beliefs – from debates over quantum theory to discussions of metaphysics – have no clear reproductive relevance at all. How, and whether, such beliefs are related to cognitive mechanisms that have been selected for veracity is uncertain (Cromer 1993; Wolpert 2000). What does not seem uncertain is that manifold beliefs do not influence behaviors or the behaviors they do influence are not reproductively salient. Belief-forming mechanisms generate variety that, analogous to neutral polymorphisms (Kimura 1991), may be unpruned by the adaptive consequences of their truth or falsity. Indeed, the capacity for some degree of cognitive licentiousness may itself be an adaptation to the "uncertain futures problem" (Plotkin 1997; Wagner 2005).

Finally, M&D's conclusion requires the falsity not only of the above ways in which selection fails to exclude misbelief, but also of the more global but controversial thesis that nothing at all about the process of natural selection serves to favor truth-conducive cognitive tools (Churchland 1987; Plantinga 2002; Stich 1990).

On selectionist accounts of the origin of mind, beliefs and belief forming mechanisms are selected by virtue of their capacity to support adaptive behavior or internal states. Thus, belief forming mechanisms will be selected when they yield (i) a representational model that orients organisms towards adaptive behaviors, and/or (ii) a correlational source of arousal or inhibition that serves to motivate adaptive (or inhibit maladaptive) behavior. The question then becomes: Are models that are true better at orienting organisms towards adaptive behaviors, or are true beliefs better at arousing effective desires for adaptive behaviors? From what we know about the action of natural selection, the most prudent answer may be: "There is no reason to think so."

Why is there no reason to think so? Because (in science, and in belief generally) models need only to "save appearances" in order to be successful. Consider the task of designing "thinking" robots for a competition in which the winners were duplicated (with minor program variations) for future competitions. While one would surely seek to program competing robots to form beliefs that provided an isomorphic "map" of the external environment, would one further seek to program beliefs about that environment that were true? Not obviously. Indeed, there are numerous ways of programming the robot to "conceptualize" its environment that, while representationally biased or even radically false, are nonetheless (a) appropriately isomorphic and (b) reliably adaptive behavior-inducing. Such programs would be adaptive.

What is true of programmed learning robots is true of selection-designed cognition. Dennett has aptly commented, "Lying behind, and distinct from, our reasons are evolutionary reasons, free-floating rationales that have been endorsed by natural selection" (Dennett 2006a, p. 93). Our reasons (in better moments) are truth-seeking; natural selection's are fitness seeking. We cannot know if, in achieving its reasons, selection allows us also to achieve ours.

Of course, one might respond that just because our belief-forming mechanisms are liable to error in these domains does not mean that they are routinely or irremediably unreliable (after all, we often discover our errors, like the cognitive biases mentioned above). But this offers little reassurance, since the seeming discovery of error relies on comparing beliefs to other beliefs which, for all we know, are comparably unreliable, though perhaps for different reasons.

Richard Dawkins has commented that "however many ways there are of being alive, it is certain that there are vastly more ways of being dead" (Dawkins 1996, p. 9). The same is true of being right and wrong. Natural selection is immensely effective at weeding out ways of not being alive. It is unclear how well it fares in culling ways of not believing truly.

Culturally transmitted misbeliefs

doi:10.1017/S0140525X09991348

Dan Sperber

Institut Jean Nicod, ENS, 75005 Paris, France.

dan.sperber@gmail.com www.dan.sperber.fr

Abstract: Most human beliefs are acquired through communication, and so are most misbeliefs. Just like the misbeliefs discussed by McKay & Dennett (M&D), culturally transmitted misbeliefs tend to result from limitations rather than malfunctions of the mechanisms that produce them, and few if any can be argued to be adaptations. However, the mechanisms involved, the contents, and the hypothetical adaptive value tend to be specific to the cultural case.

Most of humans' beliefs, or at least most of their general beliefs, are acquired through communication. I owe my beliefs that I was born in Cagnes-sur-mer, that Washington is the capital of the US, that mercury is a metal, that dodos are extinct, that stagflation is bad, and so on ad indefinitum, not to my own perceptions and inferences on those matters, but to the words of others. Are these beliefs "grounded" in McKay & Dennett's (M&D's) sense, that is, "appropriately founded on evidence and existing beliefs" (target article, sect. 1, para. 2)? Not on relevant evidence and beliefs available to me. I hold these beliefs because I trust their sources (or, anyhow, trusted them at the time I formed the beliefs). My trusting of sources may itself be founded on appropriate evidence of their trustworthiness, but quite often it is founded rather on my trust of yet other sources that have vouched for them; for instance, I trusted the textbooks I read because I trusted the teachers who vouched for them, and I trusted the teachers because I trusted my parents who vouched for them. Needless to say, the authors of the textbooks themselves were just reporting information from yet other sources.

Of course, however long the transmission chain, communicated beliefs may be vicariously grounded in appropriate

evidence and background beliefs that had been available to the initial communicators. Nevertheless, long chains of transmission carry serious epistemic risks of two kinds. First, judgments of trustworthiness are less than 100% reliable, so that, generally speaking, the longer the chain, the lesser its compounded reliability (and this even if, serendipitously, the initial source of the transmitted belief happens to be have been trustworthy). Second, information is typically transformed in the process of transmission. As a result, a belief at the end of the chain is quite often different in content from the one at the beginning and therefore cannot vicariously benefit from initial grounding. This is particularly true of orally transmitted cultural beliefs, notably religious beliefs of the kind studied by anthropologists. One generation's religious beliefs may undergo changes in its lifetime and anyhow is a transformation of the beliefs of the previous generation. There is no initial religious belief at the dawn of time, but rather, an increasing – and sometimes decreasing – religious tenor in a variety of beliefs; later beliefs are not copies of earlier ones.

The absence of appropriate grounding not just of religious beliefs, but of so many others cultural beliefs concerning, for example, food, health, or the moral traits of ethnic groups, means that human population are inhabited by a host of poorly grounded or ungrounded beliefs. Most of these are, in the terms of M&D, misbeliefs. In fact, most of our misbeliefs are culturally transmitted misbeliefs rather than individual mistakes, distortions, or delusions.

Does this mean that the social and cognitive mechanisms through which we come to hold cultural misbeliefs are malfunctioning? Are humans irrationally gullible? No, the prevalence of cultural misbeliefs is compatible with the view that the mental mechanisms involved in epistemic trust (Origgi 2004) and epistemic vigilance (Mascaro & Sperber 2009; Sperber et al., forthcoming) are calibrated to filter information in interpersonal communication, if not optimally, at least reasonably well. They do, however, create a susceptibility to misinformation that originated not in one's direct interlocutors but long before in extended chains of transmission. This vulnerability is enhanced when it is well beyond the individual's competence to assess the truth or at least the plausibility of the contents transmitted. This is particularly the case when the contents in questions are too obscure to be open to epistemic assessment.

In the process of cultural transmission and transformation, beliefs may lose not only their empirical grounding but also their epistemic evaluability. For a belief to be evaluable, it must have a propositional content, that is, be true-or-false. One may relax the criterion so as to take into account the fact that many, possibly most, of our beliefs are not sharply propositional and may, in a range of limiting cases, lack a truth value. Still, for beliefs to be informative and guide action, they had better, in most ordinary situations, be such that their relevant consequences, practical consequences in particular, can be inferred. Many culturally transmitted beliefs do not satisfy this criterion. Their content is not just vague; it is mysterious to the believers themselves and open to an endless variety of exegeses. These are what I have called semi-propositional or half-understood beliefs (Sperber 1982; 1997). The paradigmatic example of a semi-propositional belief is the dogma of the Holy Trinity, which the believers themselves insist is mysterious. Of course, philosophers who define a belief as an attitude *towards a proposition* may dispute that "semi-propositional beliefs" are beliefs at all. But from a cognitive and social science point of view, a definition of *belief* that excludes most religious beliefs renders itself irrelevant. In particular, it disposes by definitional fiat of a wide class of cultural beliefs of which it can be disputed whether they are false or lack truth value, but that are definitely not true and hence are misbeliefs (even religious believers would accept this of religious beliefs other than their own, i.e., of the vast majority of religious beliefs).

I have long argued that cultural misbeliefs occur and propagate as a by-product, a side-effect of our cognitive and communicative dispositions (Sperber 1985; 1990). Still, it could be that some of these misbeliefs or some classes of them contribute to the reproductive success of their carriers in a manner that indirectly contributes to their own propagation. One possible class of such adaptive cultural misbeliefs would be beliefs the expression of which contributes to group identities and solidarities that enhance the individual's fitness. Unlike the positive individual illusions discussed by M&D, the adaptiveness of such beliefs does not come from the manner in which their content guides the believers' actions. It is not the content of the beliefs that matters; it is who you share them with. Yet not just any content is equally appropriate to serve such an adaptive role. In particular, a content unproblematically open to epistemic evaluation might either raise objections within the relevant social group, or, on the contrary, be too easily shared beyond that group. So, semi-propositional contents are ceteris paribus better contents for beliefs the adaptive value of which has to do with cultural sharedness, not because these contents contribute to this adaptive value by guiding action, but because they do not stand in the way of acceptance by the relevant group. Their content may also have features that contribute positively to their cultural success, for instance by rendering them more memorable, but this is another story (see, e.g., Atran & Norenzayan 2004; Boyer 1994; Sperber 1985).

Adaptive misbeliefs and false memories

doi:10.1017/S0140525X09991488

John Sutton

Macquarie Centre for Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia.

jsutton@maccs.mq.edu.au

http://www.phil.mq.edu.au/staff/jsutton

Abstract: McKay & Dennett (M&D) suggest that some positive illusions are adaptive. But there is a bidirectional link between memory and positive illusions: Biased autobiographical memories filter incoming information, and self-enhancing information is preferentially attended and used to update memory. Extending M&D's approach, I ask if certain false memories might be adaptive, defending a broad view of the psychosocial functions of remembering.

Positive illusions, including those that "propel adaptive actions" (target article, sect. 13, para. 6) are maintained over time even (within limits) in the face of recalcitrant evidence. So they require sophisticated intertemporal accounting: Memory and associated forms of mental time travel must be enlisted if positive illusions are to be stable enough to enhance fitness, to be "pervasive, enduring, and systematic" rather than mere temporary errors (Taylor & Brown 1988, p. 194). So if McKay & Dennett (M&D) are right that certain kinds of ungrounded belief are adaptive, theories of memory are directly implicated. This link extends M&D's account of adaptive misbeliefs, suggesting new questions for memory research.

The sparse literature on functional analyses of remembering addresses the adaptive nature of forgetting and the puzzling luxury of autobiographical memory (Bjork & Bjork 1988; Boyer 2008a; 2009; Glenberg 1997; Nairne 2005; Nairne et al. 2007; Schacter 2001). But the possibility that false memories (or ungrounded memories, which often contingently turn out false) could themselves be adaptive is surprising. False memories are usually seen as unfortunate outcomes of the constructive nature of remembering (Bernstein & Loftus 2009, p. 373), just as the manipulability of general belief-fixation is seen as epistemological trouble. But this standard line of thought is too quick, on two counts: reconstruction is not itself always distortion