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

(Barnier et al. 2008; Sutton 2009), and, as M&D suggest, falsity need not always be maladaptive.

Biased contextual and autobiographical memories filter incoming information, and, in turn, self-enhancing (or otherwise illusory) information is preferentially attended and used to update memory. This bidirectional link between memory and positive illusions lies at the heart, for example, of temporal self-appraisal theory (Wilson & Ross 2003). Memory directly supports positive illusions: When unrealistic inflation of *current* self-image is difficult, people still "selectively recall and reconstruct evidence from the past that makes them feel good about their current selves" (Wilson & Ross 2001, p. 582). Conversely, the motivation to maintain positive self-regard in the present motivates misplaced or false memories, for example, as we subjectively move favourable past events forwards in time and unfavourable past events backward (Ross & Wilson 2002, p. 800). These loops between autobiography and the control of action, the self remembered and the working self, exhibit considerable variation (Conway 2005). Some theorists stress strong reflexive feedback from self-representation into behavior, with ongoing integration lived out between actions and self-ascribed character, emotions, memories, and plans (Velleman 2006). Others note gulfs between the story and the life, seeing narrative self-descriptions more like public relations reports that float free of the causal processes behind the government's or organization's behavior (Clark 1994; Dennett 1991): This need not be morally or psychologically suspect, due to deliberate suppression or self-deceit, for the narrative and memory capacities – just like PR spokespersons – often don't have nor need the knowledge, or the contacts, or the access, either to get the back-story right or to directly feed in to future choices and actions.

Could false memories, as M&D might suggest, in some circumstances enhance affective, cognitive, and social well-being? A broad positivity bias in autobiographical memory is linked to enhanced emotion regulation: the tendency to accentuate the positive in recalling experiences from the personal past drives improvements in mood (Mather & Carstensen 2005). Such memory biases are most securely demonstrated in older adults, whose positive affect is increased when reminiscing about positive past experiences (Pasupathi & Carstensen 2003); but younger adults get the same emotionally enhancing effects of biased autobiographical remembering when they remember while focussing specifically on their emotional state rather than on accuracy (Kennedy et al. 2004). That this positivity bias drives false memories, not just general selectivity in recall, is suggested by a recent study in which older adults recall more false *positive* memories than false negative memories across a range of stimuli (Fernandes et al. 2008).

Narcissistic biases are thus not found only in subpersonal sensory systems. Just as the "narcissistic encoding" of sensory information is driven by our "sensory-motor projects" rather than by any ontological project of correctly cataloguing the world (Akins 1996, p. 370), so the range of past experiences to which we are maximally sensitive in remembering are those relevant to embedded action, or salient for current and future decision-making. The gulf between design for fitness maximization and design for truth preservation (target article, sect. 9) may be widest in relation to truths about the past, especially about the distant past and about particular past events. As Boyer notes, the human "capacity to store unique episodes" is strange when "organisms learn about the past mostly to the extent that they can extract from past situations what is *not* unique about them, and what will be relevant in the future" (Boyer 2009, p. 4). So, given the likely recent emergence of mental time travel, its biological function may involve just the broader self-related and social functions at work in the positive illusions described by M&D. In remembering, correspondence with reality is often trumped by coherence, truth by psychosocial utility (Alea & Bluck 2003; Conway & Pleydell-Pearce 2000; Wilson & Ross 2003). Evolved tendencies for forming memories which lack perfectly secure grounds are useful computational short-cuts, minimizing the heavy cognitive costs of precise source monitoring, of tracking the diverse origins of our representations of what is not present in the immediate environment. The idea parallels the suggestion that forgetting is an adaptive response to the computational challenges of retrieval access in context-sensitive systems (Anderson & Schooler 2000; Michaelian, submitted). But M&D's approach suggests the extra thought that sometimes-false memories can themselves, given the structure of the social and natural environment, be *intelligent* errors. Not only are the costs of trusting and acting on them in general less than those of strenuously applying forensic standards of warrant and justification, but (further) misremembering things in particular positive ways might have direct personal, motivational, and social benefits. Neither social nor motivational influence on memory is intrinsically malign: M&D remind memory researchers that cataloguing our susceptibility to misinformation in non-standard, experimentally skewed environments may blind us to the richer functions of remembering.

Effective untestability and bounded rationality help in seeing religion as adaptive misbelief

doi:10.1017/S0140525X0999135X

Konrad Talmont-Kaminski

Institute of Philosophy. Marie Curie-Sklodowska University, 20-031 Lublin, Poland.

ktalmont@bacon.umcs.lublin.pl

http://bacon.umcs.lublin.pl/~ktalmont

Abstract: McKay & Dennett (M&D) look for adaptive misbeliefs that result from the normal, though fallible, functioning of human cognition. Their account can be substantially improved by the addition of two elements: (1) significance of a belief's testability for its functionality, and (2) an account of reason appropriate to understanding systemic misbelief. Together, these points show why religion probably is an adaptive misbelief.

McKay & Dennett (M&D) think that systematically adaptive misbeliefs are unstable – that over time they are revealed to be false, much as in the case of the boy who cried wolf. This need not be the case. Systematic falsehood can only be discounted or accommodated if its truth can be investigated effectively. Some claims are particularly difficult to investigate in some contexts, however. The reasons are threefold (explored in Talmont-Kaminski 2009). First, their content may be such as to impede investigation. Supernatural claims, for example, typically involve entities that are invisible, shy, or just far, far away. Second, there may be social taboos against investigating such claims. Durkheim's (1912) category of the sacred singles out precisely this kind of social barrier against investigation. Third, which claims can be investigated depends upon which scientific tools are available. Opposition to scientific development ensures that some misbeliefs remain uninvestigated. If people think the wolf is invisible, they won't be surprised when they don't see it. Similarly, if looking for it would break some social mores or would, for example, require heat-sensing cameras they cannot access, people would probably be held back from checking the boy's claims.

Without the possibility of investigation, beliefs become untethered from their truth-value so their popularity and stability are free to be determined by the idiosyncrasies of the human belief-forming system (BFS) as well as any functions that they might have – it is the *untestable* beliefs that can be most readily moulded to best serve their function. The problem, therefore, is not belief that runs counter to evidence but belief without