

changed contexts. People modify their autobiographical memories to fit implicit “theories of change.” They, for instance, think that one gets better at a particular task with practice and therefore revise their memories of past performance to fit the predicted performance curve (Ross & Wilson 2003). In a similar way, in hindsight protocols people revise memories of their own prior guesses (e.g., that London has 10 million inhabitants) after receiving feedback information. Most familiar is attitude-revision, in which subjects routinely mis-remember previously held and subsequently changed attitudes.

These distortions seem to result from the normal standard operation of memory systems. Yet they result in misbelief. Why is that the case?

Distortion is a normative notion, so what is the standard against which memory systems are failing? Surprisingly, this is generally left implicit in memory research. In contrast to, say, decision-making, in which human “biases” are described as deviations from normative models, there are no explicit standards in memory research. That is because an explicit standard for memory performance would require a description of memory functions, and traditionally memory researchers have not been overly preoccupied by functional considerations, with a few exceptions (Anderson & Schooler 2000; Nairne et al. 2008).

As a consequence, memory performance is evaluated against generally tacit, apparently self-evident commonsense assumptions – we can infer those assumptions from the very fact that some memory processes are treated as “distortions.” As mentioned above, it seems that they constitute deviations from a tacit and largely implausible view of memory systems. One assumption seems to be that memory as storage of information is not subject to the same cost-benefit constraints as the rest of cognition, so that information acquired should be stored rather than transformed, *pace* Bartlett (1932). Another assumption is that memory retrieval has its own function, independent from decision-making, so that one should, for instance, expect people to recall attitudes that did not lead to particular decisions.

But both assumptions are biologically odd. It makes obvious sense to consider memory retrieval as a biological function that comes at a cost and is therefore designed to maximize return on that cost (Dukas 1999). Also, it makes evolutionary sense to keep in mind that organisms do not develop cognitive abilities (e.g., retrieval of past experience) for abstract epistemic benefits (knowing what used to be the case). They retrieve information inasmuch as it helps fitness-enhancing decision-making in the present (Suddendorf & Corballis 2007).

Seen in this perspective, many cases of “distortion” appear highly functional. Consider misinformation and other situations in which memories are influenced by confederates’ suggestions. The possibility and need of acquiring vast information from conspecifics also creates the possibility of error and deception. For each item of information, memory and decision-making systems must, implicitly or explicitly, assess the costs and benefits of including information in a belief-box or, alternatively, of keeping track of the information’s “source-tag.” It is certainly plausible that, *in some circumstances*, it is too costly to keep the source-tags for many items of information if they are all used to build a coherent, usable account of one’s own experience. In the same way, repetition effects show that internal judgments of familiarity and fluency play an important role in decision-making. Intuitive epistemics here uses the external world regularity that *in some circumstances* true information is more frequent than false information. What matters for adaptive design is that the circumstances in question be such that this sort of decision-making does not lead to *excessive* vulnerability.

Now turn to attitude revision. In a functional perspective, accurate memory of past attitudes would be an odd proposition for a well-designed memory system. To preserve traces of past, now-irrelevant attitudes without compromising its computations, the system would need to quarantine them from on-line motivation and decision-making (Cosmides & Tooby 2000). The

## Extending the range of adaptive misbelief: Memory “distortions” as functional features

doi:10.1017/S0140525X09991397

Pascal Boyer

Department of Psychology, Washington University in St. Louis, St. Louis, MO 63130.

pboyer@artsci.wustl.edu

http://artsci.wustl.edu/~pboyer

**Abstract:** A large amount of research in cognitive psychology is focused on memory distortions, understood as deviations from various (largely implicit) standards. Many alleged distortions actually suggest a highly functional system that balances the cost of acquiring new information with the benefit of relevant, contextually appropriate decision-making. In this sense many memories may be examples of functionally adaptive misbelief.

Memory illusions or distortions are a major area of recent research (Brainerd & Reyna 2005; Roediger 1996; Schacter & Coyle 1995). They are very diverse, ranging from intrusions in word-list recall to therapy-influenced imaginings of previous lives or systematic abuse.

Dramatic memory distortions seem to influence belief-fixation. For instance, in the illusory truth effect, statements read several times are more likely rated as true than statements read only once. People who repeatedly imagine performing a particular action may end up believing they actually performed it (imagination inflation). Misinformation paradigms show that most people are vulnerable to memory revision when plausible information is implied by experimenters. In social contagion protocols, people tend to believe they actually saw what is in fact suggested by the confederate with whom they watched a video.

Another major type of distortion is revision of prior mental states under the influence of newly received information or

extra cost of such computational “cordoning off” of memories may not be offset by the advantages, if any, of maintaining a record of past attitudes. In the same way, schema-based biased reconstruction of autobiographical memories, as occurs when people hold a particular, often implicit “theory of change” for a particular domain, may also contribute to efficient here-and-now decision-making by saving costs on specific but irrelevant episodic traces (Klein et al. 2002). Finally, a hindsight bias may constitute the most efficient way of making updated information more accessible than wrong information (Hoffrage et al. 2000). In such a perspective, the study of memory “distortions” could be part of a functional account of the systems involved, as is the case for perceptual illusions (Roediger 1996).

Is all this adaptive? An evolutionary perspective on memory cannot maintain the assumption of a frictionless, cost-free recording of experience that seems to be the implicit standard in memory research. Memory need be only as “good” as the advantage in decision-making it affords, measured against the cost of its operation (Nairne et al. 2008). This is why we go around assuming that we always knew what we now know, and believed the same beliefs; and we often construe as direct experience what we only know from others’ reports – but all this is part and parcel of having a highly efficient memory system. If that is the case, it may well be that a great number of our memories, as beliefs about past occurrences, are instances of adaptive misbeliefs.