

Non-instrumental belief is largely founded on singularity¹

Commentary on McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32(6), 493-510.

by George Ainslie

151 Veterans Affairs Medical Center, Coatesville, PA 19320
University of Cape Town, Rondebosch 7701, South Africa

george.ainslie@va.gov
www.picoeconomics.org

This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA, USA. The opinions expressed are not those of the Department of Veterans Affairs of the U.S. Government. It is considered a work of the U.S. government and as such is not subject to copyright within the United States.

Abstract:

The radical evolutionary step that divides human decision-making from that of nonhumans is the ability to excite the reward process for its own sake, in imagination. Combined with hyperbolic over-valuation of the present, this ability is a potential threat to both the individual's long term survival and the natural selection of high intelligence. Human belief is intrinsically "unfounded" or under-founded, which may or may not be adaptive.

Text

McKay & Dennett (M&D) depict the category of adaptive groundless beliefs as a small, albeit fascinating, exception to their "default presumption – that true beliefs are adaptive and misbeliefs maladaptive" (target article, sect. 15, para. 3). They review many kinds of examples, such as self-confirming beliefs (placebos), beliefs that by their nature cannot be tested (faiths), and beliefs that could be tested but are not (delusions). The striking feature of these cases is that they are not sharply demarcated from grounded beliefs, and thus represent not a small cabinet of curiosities but demonstrations of a basic inadequacy in the conventional understanding of belief. The authors start toward repairing this inadequacy by pointing out that in many cases, "[beliefs'] adaptiveness is independent of their truth or falsity" (sect. 14, para. 2). This implies that beliefs are ultimately selected for functionality, but it raises the question exemplified by M&D's quote from Humphrey: "If people can get better by their own efforts, why don't they just get on with it as soon as they get sick – without having to wait, as it were, for outside permission?" (Humphrey 2004, p. 736, cited in sect. 13). The authors analyze the problem in terms of adaptiveness, but really cannot do without a key intervening variable, reward.

It is certainly true that “the driving force behind natural selection is survival and reproduction, not truth” (from Bloom [2004], quoted in sect. 15). However, evolution has developed the reward process as a proxy for survival and reproduction, outcomes that are too global and usually too distant to select the behaviors of individual organisms. Although “survival is the only hard currency of natural selection” (sect. 15. last para.), it affects choice only by backing the token currency of reward; to the extent that organisms are engineered to learn at all, they are engineered to maximize prospective reward, which is the quantity that must have an “exchange rate with truth” (sect. 15, concluding para.). A further constraint is that the valuation of prospective reward seems to be fundamentally tied to the Weber-Fechner law by which most psychophysical quantities are perceived (Gibbon 1977), causing it to be discounted for delay in a hyperbolic curve rather than a “rational,” exponential curve (Green & Myerson 2004; Kirby 1997). The exchange rate of reward with truth is necessarily close to parity when animals have only “aliefs,” not beliefs (sect. 8), and the hyperbolic over-valuation of imminent rewards should not matter when animals’ long-term interests are served by instincts that make long-term preparations such as hoarding, dam building, and migrating rewarding in the short term. However, with selection for increasing intelligence has come increasing imagination, and with imagination the unhitching of reward from adaptiveness, of short-term from long-term interests, and of belief from truth.

Imagination is governed by reward. It discovers short cuts that detach reward contingencies from the adaptive functions that originally selected for them. People have learned to mate without reproducing, fight without needing to, and commit themselves to costly hobbies that do not contribute to surviving offspring. Furthermore, we have learned to rob future welfare for present pleasure, not just with addictive substances but also with socially accepted activities ranging in excitement level from death-defying adventure to simple procrastination (Ainslie, 2010). Most important for the present discussion, we have learned to make occasions for current emotional reward from events that are not currently happening, in the form of memories, fantasies – and beliefs (as opposed to aliefs). Intelligence obviously has many adaptive features, but its ongoing evolution must be limited – probably has already been limited – by the availability of constraints on the urge for diverting long-term resources to current consumption. This is the context in which we need to address the question of why people cannot get well, or get confident, or get happy, “without having to wait ... for outside permission.”

Where reward is strongly bound to survival resources – food, warmth, avoidance of injury – the cost of misbelief will be deprivation or pain, so instrumental beliefs will be constrained mostly by their predictiveness, as the authors also note. Where hardwired sensations are not involved, or even where they are significantly delayed, the prospective benefits and costs of belief obviously depend on the reward that can be expected from imagination. Sources of this reward vary widely; for instance: sublime fantasy, puzzle-solving, vicarious adventure, or gratification of urges to obey compulsions or entertain anxiety or disgust. We learn to imagine various scenarios on various occasions based on the patterns of reward that ensue, cultivating some feelings and avoiding or resisting the urges for others. Belief might be best defined as the faculty that directs imagination so as to improve long-term outcomes, relative to the results of spontaneous immersion in the

moment; but this improvement is measured in reward, which corresponds to adaptiveness only to the extent that evolution has had time to modify the proxy function of reward to keep up with increasing *Homo* intelligence. And there remains motivational pressure for belief to serve spontaneous immersion, in the form of wishful thinking.

I have described elsewhere how hyperbolic discounting of reward predicts regularities in the competition of reward-seeking processes (Ainslie 2001, pp. 48–104, 161–97; 2005). Here the important aspect is that imagination ad lib exhausts itself in premature payoffs. When one occasion for reward is as good as another, they will replace each other randomly, and the imagining will have the quality of a daydream. Conversely, if there is a single, relatively rare occasion that stands out from the others, it will make the corresponding imagination robust. The experience of such *singularity* may be much like that of having solved a puzzle or detected a fact of nature. The occasion in question will stand out from the common ruck of imaginings just as a fact stands out from a fantasy.

Where information about the natural world is absent or ambiguous, singularity may be the best clue about how it functions – parsimony is a decent starting place for theories. But a belief that distinctly delivers good news and bad news will be productive of reward in its own right, regardless of its eventual accuracy. The emotional effectiveness of singular occasions may be experienced as a kind of factuality, more or less confounded with the factuality that comes from physical observation. In the most conspicuous cases, remembered events are experienced again on their anniversaries, especially when the anniversary is a round number; original works of art are felt to be more “real” than exact copies; and placebos (as in sect. 13) are effective in proportion to the expensiveness of the ingredients or the prestige of the healer. Even realistic beliefs get additional value by serving as occasions for emotional reward, as in the “drug effect” of money (Lea & Webley 2006). Conversely, faced with unwelcome urges such as hypochondria, phobic anxiety, or a sense of being dirty, a person searches for a favorable interpretation of the situation – whether she can feel well, or safe, or clean. This interpretation cannot be arbitrary; wishes have little impact. She must choose her belief on the basis of “facts” that she discerns in events beyond her control – a pill given by a doctor, a lucky charm or safety signal, or a “scientific” disinfectant. The belief may even become stabilized as a personal rule: in effect, “I will not give in to panic or disgust when this signal is present.” The same role of singularity can be seen in many other misbeliefs. For instance, delusions (sect. 9) tend to be based on a logical deduction or a remarkable coincidence, and religious faiths (sect. 11) depend on the singularity that comes from having had long histories of consensual agreement – hence their fear of heresies. It would be fruitless to try to decide whether such hedonically based beliefs are more or less adaptive than veridicality; evolution veered away from veridicality with the apes.

1. The author of this commentary is employed by a government agency, and as such this commentary is considered a work of the U. S. government and not subject to copyright within the United States.

References

- Ainslie, G. (2001) *Breakdown of will*. Cambridge University Press.
- Ainslie, G. (2005) Précis of *Breakdown of will*. *Behavioral and Brain Sciences* 28(5):635–73
- Ainslie, G. (2010) Procrastination, the basic impulse. In: *The thief of time: Philosophical essays on procrastination*, ed. C. Andreou & M. White. Oxford University Press, pp. 11-27
- Bloom, P. (2004) *Descartes' baby: How child development explains what makes us human*. Arrow Books.
- Gibbon, J. (1977) Scalar expectancy theory and Weber's law in animal timing. *Psychological Review* 84:279–325
- Green, L. & Myerson, J. (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin* 130:769–92.
- Humphrey, N. (2004) The placebo effect. In: *Oxford companion to the mind*, 2nd edition, ed. R. L. Gregory, pp. 735–36. Oxford University Press. Available at: www.humphrey.org.uk
- Kirby, K. N. (1997) Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General* 126:54–70.
- Lea, S. E.G. & Webley, P. (2006) Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences* 29:161–209.