**Title: Attack of the Memes! How cultural parasites can subvert human interests**

**Authors**

Maarten Boudry (corresponding author)

Ghent University - Department of Philosophy & Moral Sciences

St.-Pietersnieuwstraat 49 - room 204

9000 Ghent - Belgium

+ 32 473 92 38 75

maartenboudry@gmail.com

Steije Hofhuis

Utrecht University - Department of History and Art History

Drift 6

3512 BS Utrecht, The Netherlands

[s.t.hofhuis@uu.nl](mailto:s.t.hofhuis@uu.nl)

Table of contents

[Introduction 3](#_Toc466804196)

[Cultural evolution 4](#_Toc466804197)

[Human authorship 4](#_Toc466804198)

[What’s in a meme? 5](#_Toc466804199)

[Millikan’s challenge 7](#_Toc466804200)

[Failed answers to Millikan’s challenge 8](#_Toc466804201)

[1. Panmemetics 8](#_Toc466804202)

[2. Biologically maladaptive culture 10](#_Toc466804203)

[Doxastic parasites 12](#_Toc466804204)

[Explaining the appeal of misbelief 12](#_Toc466804205)

[Doxastic Catch-22 14](#_Toc466804206)

[Case study: European witch persecutions 15](#_Toc466804207)

[Discussion 17](#_Toc466804208)

**Attack of the Memes! How cultural parasites can subvert human interests**

**Abstract**

Are there any such things as *mind viruses*? By analogy with biological parasites, such cultural items supposed to subvert or harm the interests of their host. Most popularly, this notion has been associated with Richard Dawkins’ concept of the “selfish meme”. To unpack this claim, we first clear some conceptual ground around the notion of cultural adaptation and units called ‘memes’. We then formulate *Millikan’s challenge*: how can cultural items (‘memes’ or whatever you want to call them) develop novel purposes of their own, cross-cutting or subverting human purposes? If this central challenge is not met, meme talk will be vacuous or superfluous. First, we discuss why other attempts to answer Millikan’s challenge have failed. In particular, we put to rest the claims of *panmemetics*, a somewhat sinister worldview which treats all of culture as swarms of selfish agents. Next, we reject a more reasonable, but still overly permissive approach to cultural parasitism, which equates mind parasites with *biologically maladaptive* culture. Finally, we present our own answer to Millikan’s challenge: certain *systems of misbelief* can be fruitfully treated as selfish agents developing novel purposes of their own. Such mind parasites are designed to spread in a viral-like manner, without any regard to the interests of their human hosts. As a case study, we discuss the witch hunts of early modern Europe. In this particular case, adopting the meme’s eye view promises to shed new light on a mystery that historians and social scientists have been wrestling with for decades.

**Keywords**: meme’s eye view; mind parasites; cultural adaptation; misbeliefs; witch persecutions; maladaptive culture

# Introduction

Few ideas are so well applicable to itself, as even its critics might admit, as the notion of a *meme*. Introduced by Richard Dawkins as a unit of cultural information, a rival replicator to the gene, memes have captured the public imagination, and colonized brains far and wide. But is it a valuable idea, which *deserves* to be spread? Or is it a vacuous and worthless gambit for churning out pseudo-explanations of culture, a misleading metaphor that has muddled the minds of many enthusiasts and has even fostered a syndrome called “Darwinian paranoia” ([Godfrey-Smith, 2009](#_ENREF_28))? In short, is the *meme* meme itself an annoying mind parasite?

Meme enthusiasts have made some ambitious claims on behalf of memes. The science of memetics, as it came to be known, promised to offer a new and unifying theory of culture, superseding traditional conceptions about human agency, intentionality and free will ([Aunger, 2002](#_ENREF_3); [Blackmore, 2000](#_ENREF_8); [Lynch, 2008](#_ENREF_42)). In one slogan, the startling thesis of memetics is: *we think we are in control of our thoughts, but they are in control of us*. Such overblown pretensions, to put it mildly, have not enhanced the fitness of the meme meme in the academic world, and have made it easy for critics to make short shrift of the whole endeavor. In addition, two implausible notions came to be associated with memes: the idea that culture is composed of discrete units, and that they spread through high-fidelity replication. In this paper, we argue that these contentious points have unfortunately obscured the real value of memes. There is no need for a new *science of memetics*, in the sense of a unifying and overarching theory of culture, in close analogy with genetics. But we do need the *meme’s eye view* ([Dennett, 1995](#_ENREF_21), [2006](#_ENREF_23)) to understand the phenomenon of mind parasites, “selfish” forms of culture that are designed to further their own propagation, while subverting the interests of their hosts. Adopting the meme’s eye view neither makes you “mind-blind” ([Atran, 2002](#_ENREF_2); [Boyer, 1994](#_ENREF_13); [Sperber, 2000](#_ENREF_58)), nor is it wedded to the thesis of simple replication or particularism. On the contrary, it draws our attention to patterns that are often missed in traditional approaches to culture, even in sophisticated theories of cultural evolution ([Henrich, 2015](#_ENREF_35); [Norenzayan et al., 2014](#_ENREF_49)).

Most scholars of cultural evolution have ignored the *meme’s eye view*, and those who have taken it seriously have often been too careless or permissive in applying it. Because the metaphor of selfish memes is so vivid and captivating, it is easy to get carried away. With a bit of creativity, one can see memes plotting and scheming everywhere. To clear the ground, we dismiss the claims of *panmemetics*, and clarify the role of human agency in replicator talk, drawing instructive analogies with artificial selection in the living world. Next, we clear up some confusion regarding the concept of maladaptive/viral/rogue culture, arguing that we should dissociate ‘biologically maladaptive’ (harmful to biological fitness) from ‘deleterious’ (harmful to individual agents). We take on specific challenges raised by some of the more thoughtful critics of memetics (Millikan, Pinker, Lewens, Sperber). Most centrally, we deal with Millikan’s challenge: how can memes introduce novel purposes in the world, cross-cutting our own? And how does this relate to human agency and intentionality? In the end, we argue that the selfish meme metaphor gains traction only when we are dealing with genuine mind viruses, parasitical forms of culture that undermine the interests of their hosts. After discussing two overly permissive answers to Millikan’s challenge (panmemetics and biologically maladaptive culture) we then present our own, more stringent approach. In particular, we argue that certain systems of misbelief (*doxastic parasites*)evolve to subvert the interests of their hosts, in light of a central conceptual paradox we call the Doxastic *Catch-22*: it is conceptually impossible to fashion one’s own misbeliefs in accordance with one’s best interests. In the evolution of misbeliefs, a novel dynamic is set in motion, which disregards human interests. To understand the rationale of systems of misbeliefs, we have to adopt the meme’s eye view. We illustrate these points with the example of European witch-hunts, which remains an unresolved theoretical puzzle to this day.

# Cultural Evolution

## Human Authorship

Nobody disputes that culture evolves, in the sense that it changes over time and forms something resembling lineages. But does that mean that Darwinian explanations, in terms of variation and selection, can be fruitfully applied to the cultural domain? Many scholars have argued that, even though cultural systems obviously display functional complexity, the origin of all that design work is hardly a mystery: we did it. Biological adaptations are the result of blind evolution, but cultural adaptations are the products of intelligent design. Cultural design emerges, as Pinker ([1997, p. 209](#_ENREF_51)) wrote, when someone “knuckles down, racks his brain, musters his ingenuity, and composes or writes or paints or invents something”.

However, researchers in the burgeoning field of cultural evolution have convincingly argued that plenty of adaptive design in cultural systems is “unauthored”, in the sense that it cannot be credited to the foresighted work of specific intelligent designers ([Boyd & Richerson, 2005](#_ENREF_12); [Henrich, 2015](#_ENREF_35); [Mesoudi, Whiten, & Laland, 2006](#_ENREF_44); [Richerson & Boyd, 2005](#_ENREF_54); [Wilson, 2003](#_ENREF_67)). It is now widely accepted, at least by scholars of cultural evolution, that much adaptive cultural design was not purposefully crafted by a single author, or even a few innovators and pioneers, but slowly accumulated over time, through a relatively blind process of trial and error. No single person, or even a small group of innovators, can take credit for the design work that goes into igloos, canoes, folklore tales, folk medicine, marriage institutions, procedures for detoxifying food, or complex religious rituals. Much of the design credit, as was argued most forcefully by Henrich ([2015](#_ENREF_35)), should be assigned to blind cultural evolution. Even after such cultural design has evolved, its rationale remains largely to opaque to human beings. People have little understanding of why they hunt or cook food in a particular way. When asked, they tend to give spurious *post hoc* explanations, or they draw a blank, saying in effect “this is the way we do things”. In some cases, they are not even aware of the existence of cultural design ([Henrich, 2015](#_ENREF_35)).

All this is not to say, of course, that human intelligence has no share in adaptive cultural design. But just because intelligent creatures are somehow involved in the process, as the producers and transmitters and consumers of culture, does not mean that they deserve credit for all the design work. Human beings may have local and parochial intentions, without understanding the aggregated effects of their actions over time. They may copy the most efficient spears and igloos, or imitate the recipes and habits of a successful individual, without having any clue as to the reasons for their success. They may not even be aware that they are transmitting cultural information at all. To assume that evolutionary explanations are moot wherever intentional agents are on the scene, as Pagel ([2006, p. 360](#_ENREF_50)) writes, is to “conflate intentionality with omniscience”.

In the living world, *all* the heavy lifting in design space is done by random variation and selection, or at least by some other blind and unguided processes. At least until very recently, none is the result of intelligent design. In the cultural domain, by contrast, intelligent design has varying degrees of contribution in the emergence of functional complexity. There is a continuum from deliberate, foresighted and conscious design, to mindless, unconscious and blind selection ([Dennett & McKay, 2006](#_ENREF_24)). In many cases, of course, the origins of cultural design are lost in the mist of time, leaving us with scant circumstantial evidence. Some cultural inventions were probably created in one fell swoop by geniuses, some arose in relatively isolated creative outbursts by a few pioneers, still others accumulated in small, incremental steps, with little or no understanding on the part of human agents. Sometimes the development of cultural design over time is monitored and steered in a particular direction, but sometimes it happens in a relatively unguided and haphazard way.

## What’s in a Meme?

Richerson and Boyd define culture as “information capable of affecting individuals’ behavior that they acquire from other members of their species through teaching, imitation, and other forms of social transmission.” ([Richerson & Boyd, 2005, p. 5](#_ENREF_54)) What’s in a meme then? If a piece of information with appreciable functional coherence spreads through a population in this way, with the different instantiations forming lineages of descent, then that piece of information can be baptized as a “meme”. Memes are just functional units of cultural information: beliefs, practices, rituals, words, recipes, tunes, mannerisms, pictures. Many cultural theorists resist this nomenclature, preferring instead to talk about representations or cultural variants, or the more specific terms from the above list ([Richerson & Boyd, 2005](#_ENREF_54); [Sperber, 2000](#_ENREF_58)). In their view, the term “meme” suggests that culture is composed of discrete units analogous to genes, which are copied by simple imitation. At bottom, however, this is a terminological discussion. Cultural evolutionists could decide to adopt Dawkins’ terminology, without taking on board the notion of replication by simple copying ([Durham, 1991](#_ENREF_25)). The concept of a meme, as Mesoudi et al. have argued, tolerates ambiguity and fuzzy boundaries, and does not require strict particularism. It is not a problem that we have no method for individuating one meme from another. Many of the putative objections against the concept of a “meme” would apply with equal force to the evolutionary notion of a “gene”, which is far less crisp and clear-cut than many critics imagine. In spite of this difficulty to define genes, “both empirical and theoretical traditions within population biology have thrived … by using simple, discrete gene concepts.” ([Mesoudi et al., 2006, p. 343](#_ENREF_44)) The same point applies to memes. As Richerson and Boyd write, “The selfish meme effect is quite robust. Nothing in the argument depends on cultural variants being discrete, genelike particles. It works exactly the same if “memes” were continuously varying […]” ([Richerson & Boyd, 2005, p. 154](#_ENREF_54))

For similar reasons, it is important to sidestep fruitless debates over the ontological status or physical substrate of a meme ([Aunger, 2002](#_ENREF_3), [2000](#_ENREF_4); [Lynch, 2008](#_ENREF_42)). The controversy over whether memes should be identified with particular states of the brain, or with observable artefacts or behavior, is a distraction ([Plotkin, 2002](#_ENREF_53)). Memes are most usefully thought of as pieces of information, just as genes should not be identified with DNA molecules, but with the information carried by these molecules ([Durham, 1991](#_ENREF_25); [Haig, 2007](#_ENREF_31); [Williams, 1992](#_ENREF_66)). If one cultural item is causally responsible for the emergence of another, and both are sufficiently similar to be treated as instantiating the same information, and many of these items form lineages in a human population, then it doesn’t matter whether we give them a generic title like “meme” ([Dawkins, 1976](#_ENREF_16)), “culturgen” ([Lumsden & Wilson, 1981](#_ENREF_41)) or “cultural variant” ([Richerson & Boyd, 2005](#_ENREF_54)), or if we use more traditional labels such as ideas, habits, beliefs, or artefacts. In any event, the problem of particularism is not solved by abandoning the term “meme” and substituting a different concept. When Dan Sperber, one of the foremost critic of memetics ([Sperber, 2000](#_ENREF_58)), speaks of his preferred concept of “shared representation”, he has to add essentially the same disclaimer: “what we mean is that these individuals have mental representations similar enough to be considered versions of one another.” ([Sperber, 1996, p. 82](#_ENREF_57)). Indeed. The concept of a “meme” requires nothing more than Sperber’s proviso.

The truly novel (and controversial) aspect of memetics is not its supposed particularism, nor even its putative commitment to simple replication through imitation ([Sperber, 2000](#_ENREF_58)), but the idea that memes can be treated as somehow *selfish*, as parasites of the mind. The real bone of contention is not one about terminology, but about the *meme’s point of view*: the idea that items of cultural information – whatever you want to call them – can evolve in such a way that they end up subverting the interests of their hosts. Dennett does not tire of raising this fundamental question whenever we are dealing with adaptive design through selective processes: *cui bono?* Who or what is the ultimate beneficiary of all this selecting? ([Dennett, 1990](#_ENREF_20), [1995](#_ENREF_21), [2001](#_ENREF_22))

In evolutionary biology, genes are often assigned the protagonist role in explanations of adaptive design ([Dawkins, 1982](#_ENREF_17); [Sterelny & Kitcher, 1988](#_ENREF_63); [Williams, 1992](#_ENREF_66)). By treating genes as if they were selfish agents furthering their own propagation, we can shed light on many puzzling phenomena in the living world. This is not to say that everything in evolution revolves around genes ([Jablonka & Lamb, 2005](#_ENREF_37)), or that the perspective of the organism is never useful. Organism-centered explanation of adaptive design work to the extent that their interests align with those of their hereditary material, as indeed they often do. As Dennett writes: “The theory becomes interesting only when we look at the exceptions, the circumstances under which there is a pulling apart of the two perspectives” ([Dennett, 1990, p. 130](#_ENREF_20)). In biology, there are plenty of conditions in which the perspective of the organism and the gene are pulled apart, such as kin selection, eusociality, retroviruses, parasite/host interactions, mitochondrial DNA, sex-linked genes, and transmissible cancer.

## Millikan’s Challenge

But where is the divergence of interests in the case of cultural evolution? You can baptize any piece of culture as a “meme”, but that does not mean it is usefully described as having *selfish* purposes of its own. Haig writes that memes can have purposes “to the extent that they have properties that have promoted their propagation from mind to mind.” ([Haig, 2006, p. 14](#_ENREF_30)) But are these *novel* purposes, which potentially diverge from human ones? In what sense can pieces of cultural information develop purposes that thwart the interests of their hosts? Perhaps the most lucid formulation of what we take to be the central challenge for the *meme’s eye view* was offered by Ruth Millikan:

Part of what they have been selected for is their ability to be reproduced accurately through the medium of human minds. But this does not subvert their essentially human purposes. ... The memes have merely fed these interests a much richer diet than if each person had to invent all of his own amusements, or invent all of the entertainments he uses to invoke the gratitude and appreciation of others. … Side effects and mishaps resulting from use of these [basic cognitive ] mechanisms will surely occur, but there is no reason to suppose that they systematically produce memes with purposes of a different kind from those either of the genes or of the psyche. ([Millikan, 2004, pp. 18-19](#_ENREF_45))

It is not enough just to define “memetic purposes” into existence, as long as these are merely derivative of human purposes. For the idea of selfish memes to gain traction, we need to see a divergence of interests between human and memetic purposes. Note that Millikan concedes that there can be cultural design without a designer, which was the point that troubled Pinker. She also does not object to the nomenclature of memes. Her point is that, even if memes are adapted by cultural evolution to thrive in a population of humans brains, they are still essentially serving human purposes. In order to be reproductively successful, memes have to adapt themselves to the environment of human desires, interests and tastes. The hosts may be oblivious to this selection process, but the resulting memes will still benefit them. Folksongs, recipes, artefacts, jokes and rituals are products of blind cultural evolution without ever subverting our human interests. If they can be said to go “viral”, that is only because they appeal to universal human tastes and preferences. Evolution is smarter than we are, as Orgel’s second rule reminds us: it knows better how to tickle our fancy and satisfy our cravings than any lone human genius.

If this line of reasoning is correct, it seems that memes will always serve some human interest, and it becomes vacuous to track the “fitness” of the memes themselves, as Lewens ([2015, p. 31](#_ENREF_40)) wrote. In order for the selfish meme metaphor to gain traction, however, memes must give rise to purposes of their own, which cannot be anchored in the intentions of human beings. Preferably, we would like to see cases where they actively thwart human interests, so that no other intentional explanations on the level of human agency are on offer.

Millikan’s challenge can be strengthened by a point that was made by Durham ([1991, p. 198](#_ENREF_25)). We are not just looking for memes that harm their hosts, because this may happen simply as a result of a divergence of *human* interests. It may be in the interest of one agent (or group of agents) to impose some meme on others, even though it is harmful to the latter. For example, people can be coerced into participating in a religious ritual, or adopting a marriage custom, or using a particular piece of technology, even though they dislike or disapprove of these respective memes, and even though are aware of their harmful effects. Such memes, strictly speaking, subvert the interests of their hosts, but they are still serving someone *else’s* interests. Durham calls this “selection by imposition”, as opposed to the more regular “selection by choice”, whereby individuals evaluate and select the memes that they prefer. In both cases, according to Durham, we seem to have cultural evolution through human decision making: “either way, culture changes under *human* direction” ([Durham, 1991, p. 198](#_ENREF_25)). This serves to reinforce Millikan’s point: human preferences direct cultural evolution. So how can memes have selfish interests of their own, which are not reducible to human preference?

# Failed Answers to Millikan’s Challenge

## Panmemetics

Before we discuss what we take to be *genuine* selfish memes, we have to deal with an influential and more permissive approach to meme talk, and show why it runs afoul of Millikan’s challenge. In this radical and encompassing view, which we call *panmemetics*, *all* of human culture is treated as consisting of selfish memes. Eventually, this approach leads to the subversive and startling conclusion that we are not in control of our thoughts, but the memes are in control of us. Our brains are just breeding grounds for selfish parasites manipulating us into fostering their own proliferation.[[1]](#footnote-2) In *The Meme Machine*, Susan Blackmore writes:

To start to think memetically we have to make a giant flip in our minds just as biologists had to do when taking on the idea of the selfish gene. Instead of thinking of our ideas as our own creations, and as working for us, we have to think of them as autonomous selfish memes, working only to get themselves copied. … This is a scary idea indeed. ([Blackmore, 2000, pp. 7-8](#_ENREF_8))

Another major proponents of memetics, Robert Aunger, similarly writes:

Who’s talking when I speak: the memes or me? Are my very thoughts something I was able to decide on, or are they just parasites attempting to get out of me and thus infect others? ([Aunger, 2002loc. 281](#_ENREF_3))

Elsewhere, Aunger writes about the prospect of “people turned into zombies, with only the illusion of control over their own behavior” and provocatively asks: “Do we have thoughts, or do they have us?” ([Aunger, 2002loc. 120](#_ENREF_3)).[[2]](#footnote-3) So what remains of human autonomy in this sinister worldview? Some proponents of panmemetics have offered a shimmer of hope for humanity, by exhorting us to revolt against our selfish replicators. In *The Robot’s Rebellion*, Keith Stanovich writes: “We indeed are the runaway robot of science fiction stories-the robot who subordinates its creator’s interests to its own interests.” ([Stanovich, 2005, p. xii](#_ENREF_60)). For memeticists like Susan Blackmore, however, this is “all a cop-out”. In her meme-infested worldview, there is no room for a self, for deliberate choice, for free will, so “there is no one to rebel” ([Blackmore, 2000, p. 246](#_ENREF_8)). “I” am just an illusion created by “my” memes.

If anything merits the label “Darwinian paranoia” ([Godfrey-Smith, 2009, pp. 142-145](#_ENREF_28)), this might be it. Let’s pursue the call for rebellion against selfish replicators for a second. Rebelling against that other replicator, our old friend the gene, is pretty straightforward for intelligent human beings: make sure to stay celibate or always to use contraceptives, and never visit sperm banks (or donate eggs). By adhering to these simple rules, you will doom the fate of your genes. By contrast, the notion of a “rebellion against our memes” ends up in a conceptual mire. What else is this rebellion, if not itself a “meme”, with which you might be infected upon opening Stanovich’s book? Even your critical thinking skills, which you rely on to examine the notion of selfish memes – or to evaluate the cogency of a rebellion against them – can be treated as one of those very memes that you should be rebelling against. Stanovich recognizes this fact: “Scientific and rational thinking are themselves memeplexes – co-adapted sets of interlocking memes” ([Stanovich, 2005, p. 180](#_ENREF_60)). He calls this the “co-adapted meme paradox” ([Stanovich, 2005, p. 180](#_ENREF_60)), which testifies to the “devilish recursiveness” of memetics, but in fact, it is merely an artifact of a metaphor stretched beyond the breaking point.

If selfish memes explain everything, they explain nothing. Almost any cultural phenomenon can be restaged as a little drama where the memes are pulling the strings. For example, a mundane observation such as “this apple cake recipe is popular” can be translated as “the recipe meme has succeeded in replicating itself by devising clever adaptations to appeal to human taste”. Or take the discovery of the Higgs boson: one could argue that the Higgs meme “infected” hosts all around the world following the experiments at CERN, which caused a sudden spike in its cultural fitness. From an epidemiological point of view, nothing is more contagious than the naked truth. If there is an 9.0 earthquake tomorrow in the Philippines, the attendant belief “there has been a huge earthquake in the Philippines” will infect millions of brains worldwide in a matter of hours. No biological virus can spread with such breathtaking speed. As Douglas Adams wrote: “Nothing travels faster than the speed of light with the possible exception of bad news, which obeys its own special laws.” But can anything be learned from adopting the meme’s eye view here? Is it enlightening to say that the earthquake meme is exploiting the brains of its hosts to make more copies of itself? Either this is just a fanciful and unenlightening re-description of what we knew all along, or it is a crude form of reductionism, which ignores the fact alluded to above: there is a *continuum* in culture between relatively mindless processes and intelligent design.

*Methodical selection & biotechnology*

It is instructive to compare this with the selfish gene. Because blind evolution has done almost *all* the heavy lifting in nature, and because DNA is the universal code for all life on the planet, the *gene’s eye view* is much more unifying in the living world than the meme’s eye view in the cultural world. But even the *gene’s eye view* runs up against certain limits. Methodical artificial selection, and modern biotechnology, provide interesting exceptions. Poodles, *bintje* potatoes and Belgian Blue cows are forms of biological design, accumulated through small, incremental improvements, but unlike in the case of natural selection, their lineages have been stewarded and directed by intelligent breeders.[[3]](#footnote-4) Poodles, of course, are vehicles for genetic information no less than undomesticated wolves. Their DNA gets replicated and reshuffled through meiosis every generation, just as with wolves. Can we call the poodle genes “selfish” in the interesting, metaphorical sense? In effect, from the gene’s point of view, the difference is not consequential. In the wild, wolf genes have to deal with a host of selective pressures, whereas poodle genes mostly have to deal with one overwhelming selective factor, namely the breeder’s fancy. While the breeder provides food, comfort and protection, thus slackening many of the other selective pressures, he carefully decides which poodles are allowed to mate and reproduce. Note that this is not an exception to, but merely a special case of, natural selection ([Dennett, 2001](#_ENREF_22)). One can still recast what is happening from the poodle genes’ point of view. For example, one could say that poodle genes have invaded a new niche (the bourgeois urban household), developing novel strategies for adapting to the novel selection pressures (the breeder’s fancy). Poodle genes for fluffy fur manipulate the breeder into providing care and protection, much like the brain of a cuckoo chick’s host parents are being manipulated by the parasitic bird’s bright red gape and loud begging call.

In cases of purely methodical selection, however, no-one bothers offering such a gene-centered description. Not because it is strictly false, but because it is unenlightening and contrived. Intentional explanations for poodle design are anchored in the breeders themselves, who are steering the evolution of their stock in a pre-established direction. Modern biotechnology puts further strain on the gene’s eye view. While traditional breeders are still mimicking evolution, in effect partly replacing Mother Nature’s role, bio-engineers are directly intervening in the genome to create a desired phenotype. For example, they insert a gene in cotton to allow it to make its own insecticide (Bt-cotton), or they equip the tomato genome with an anti-freeze gene. Replicator machinery is still churning along in Bt-cotton cells, but the perspective of the selfish gene falls apart.

In short, where human intentions reign and provide direction, talk of plotting replicators becomes strained, unenlightening or even preposterous. Once we realize this point, the dramatic claims of panmemetics can be put to rest.

Panmemetics is further compounded by the fact that it dissolves the distinction between good and bad forms of culture. By treating all of culture as consisting of selfish memes, furthering their own interests, we can no longer distinguish between genuine mind parasites and memes that are valuable, true or beautiful. If all memes are mind parasites of sorts, then the theory becomes vacuous ([Lewens, 2015chapter 2.5](#_ENREF_40)). If such a normative distinction appears at all in the panmemetic literature, it often seems arbitrary and bereft of theoretical significance. For example, Aunger starts his book with an interesting vignette about how witchcraft beliefs spread like viruses (see further). But it is unclear how Aunger’s panmemetic perspective helps explaining *why* witchcraft beliefs are virulent. If all memes are mind parasites of sorts, then what is the difference between witchcraft memes and, say, heliocentrism memes? The thesis that some memes are mind parasites needs to amount to more than saying that “these are some memes we don’t like” or “these are some memes that have spread far and wide”. In the absence of a substantive theoretical distinction, parlance of mind parasites reduces to mere rhetoric.[[4]](#footnote-5) We may find that antisemitism is a pernicious meme, but if this claim amounts to nothing more than rephrasing our distaste of anti-Semitism, it will indeed clarify little.[[5]](#footnote-6) To solve this problem, we will have to show that, in a substantial and non-question begging way, mind parasites are detrimental to the interests of the host *themselves*.

## Biologically Maladaptive Culture

In panmemetics, all of human culture is analyzed in terms of memes furthering their selfish interests, with human agency taking the back seat. But this approach does not answer Millikan’s challenge: in many cases panmemetics does not provide a demonstration of novel memetic purposes, but just a fanciful rephrasing of phenomena we already understood perfectly well in terms of human desires and preferences. We need a robust account of how memes introduce novel purposes in the world, and show when and where these purposes might diverge from human ones. One natural approach do to so is inspired by the analogy with biological viruses. Dennett has proposed a classification of memes analogous to biological symbionts: some are mutualists (enhancing the fitness of the host), others are commensals (neutral to the host) and still others are parasites (fitness-reducing). Biological symbionts can be placed along this continuum, and so, Dennett claims, can memes. Dennett’s classification is useful, but it brings to light an important ambiguity. In the case of biological viruses, the reference point of our classification is the fitness of the host. Properly speaking, we are classifying the different ways in which the selfish genes of host and symbiont may be related: are they antagonistic, neutral, of cooperative? But in the case of memes, their environment provides an additional and more important source of intentionality: human interests and purposes. Remember Millikan’s challenge, which was to show how memes can give rise to irreducibly novel purposes, over and above human ones. If some memes can be thought of as harmful “parasites”, as in Dennett’s classification, whose interests are they supposed to hurt exactly? Are they parasitizing on us, or on our genes? It is important to clear up this ambiguity.

Take the idea of celibacy. Delius ([1991](#_ENREF_19)) argues that “[c]elibacy is an obvious parasite meme that causes a reduction of host reproduction.” Dennett, too, writes that celibacy is “the most obvious meme example” of a parasite ([Dennett, 1995, p. 367](#_ENREF_21)). Indeed, we can treat the meme for celibacy as a “parasite” of sorts, but only when we adopt the vantage point of our genes and their desire for immortality. Any cultural invention or practice that stands in the way of biological reproduction is disastrous for human genes. On the face of it, however, the meme for celibacy need not subvert human purposes, and may precisely be the result of deliberate human choice. Even though a celibate lifestyle is often part of an evolved cultural tradition, like the celibacy rule for Catholic priests, a well-informed decision to remain childless may spring up outside of such a tradition, and would not be any less fatal for the genes. This is well demonstrated in Richerson and Boyd’s chapter on maladaptive culture in *Not by genes alone*. Richerson and Boyd give a fascinating account of the demographic transition, the phenomenon of rapidly dropping birth rates observed in many contemporary societies. Foremost among the causes of the demographic transition, they argue, are “maladaptive” forms of culture such as universal education and materialism, which spread in a population as a result of the runaway effects of prestige-biased transmission. They too, confusingly, use the term “selfish cultural variants” to describe the causes of the demographic transition ([2005, p. 153](#_ENREF_54)). But it is strange to describe people who lead voluntarily childless but fulfilling lives, as having succumbed to a selfish meme. Such a definition would not have any bearing on our personal valuation of memes. What is maladaptive from the point of view of my genes may further my personal goals in life. Richerson and Boyd seem to be aware of this tension, as witnessed by their quip, “If you want to improve your kids’ genetic fitness, for goodness sake don’t help them with their homework!” ([Richerson & Boyd, 2005, p. 178](#_ENREF_54)). Obviously, Richerson and Boyd would not dream of heeding that advice in their personal lives, nor would any other sensible parent except for the crudest pop-sociobiologist. Helping our children with their homework is exactly what *we* want, even after reflecting on the likely fitness-reducing effects. Though our preferences are themselves shaped by cultural evolution, as both Dennett and Richerson and Boyd point out ([Dennett, 1995, pp. 329-330](#_ENREF_21)), it does not seem useful to describe the memes for education as “selfish”, given that they are actively fostered and promoted by their conscious vectors.

Remember Millikan’s challenge: the memes for celibacy or contraceptives or universal education thwart our selfish genes, but they have not introduced any novel purpose in the world. The purposes of contraceptives are *our* purposes – they are explicitly represented by human producers and consumers. For example, condoms provide shortcuts to the carrot which Mother Nature dangles in front of our noses, all the while thwarting the goals she had in mind. Though it is true that a condom is itself a piece of cultural technology, pruned and perfected by the efforts of many minds, its purpose of preventing insemination is not a freely-floating memetic one, but has always been firmly anchored in the minds of conscious agents. My own reflective goals (e.g. having sex without spawning offspring) may be perfectly maladaptive in their own right, whether or not I use evolved cultural technology to achieve them. Child adoption, contraceptives, celibacy and exclusive homosexuality are all equally “parasitic” from the gene’s point of view, but in the words of Steven Pinker, himself voluntarily childless, “if my genes don't like it, they can go jump in the lake” ([Pinker, 1997, p. 52](#_ENREF_51)). In this paper, we are looking for a divergence of interests between me and my memes, not between my memes and my genes.

The difference between the two normative standards may be illustrated with one of Richerson and Boyd’s own examples: the pro-natalist ideologies of the Hutterites and Amish. These closed-knit and insular communities have succeeded in stemming the tide of what Richerson and Boyd call “maladaptive” modern culture, and have so far resisted the demographic transition observed everywhere around them. In Richerson & Boyd’s framework, Hutterite ideology inoculates its hosts against the mind parasites of modernity – such as consumerism and materialism and the idea of women’s rights – to which the rest of us have succumbed. But perhaps it is the other way around: perhaps the Hutterite belief system is the mind parasite itself, rather than the defense against it. After all, Hutterite memes may well benefit the genetic interests of their hosts, but they do not necessarily make for fulfilling and happy lives (see further).

If we are to retain Dennett’s classification between mutualists, commensals and parasites, therefore, we think it is better to define those categories with respect to human interests, not to genetic interests. There may be contexts where it is useful to contrast the interests of genes and memes, but we should not lose track of the intermediate (and more important) perspective of human agency. Memes that are “parasitic” from the perspective of my genes may simply be the outcome of deliberate human choices. Talk of “selfish memes” and “rogue culture” can be misleading here, as if humans are the hapless victims of the ideas of modernity. These are not novel memetic purposes, but distinctively human ones.

Finally, now that we have laid to rest the inflated claims of panmemetics, sidestepped the question of biological maladaptiveness, and disentangled genetic from personal interests, it is time to present what we consider *Exhibit A* of the meme’s eye view.

# Doxastic parasites

## Explaining the Appeal of Misbelief

“Like computer viruses, successful mind viruses will tend to be hard for their victims to detect. If you are the victim of one, the chances are that you won't know it, and may even vigorously deny it.” Richard Dawkins, *Viruses of the Mind* ([1993, p. 20](#_ENREF_18))

There is a sense in which doxastic memes (i.e. beliefs) are much more “infectious” than other types of culture. This is because they *command* assent. Being exposed to the right sort of evidence for belief X suffices for you to be become a carrier of X, like in the earthquake example above. It is a conceptual truth about beliefs that we do not choose our own ([Adler, 2002](#_ENREF_1)). Rather, *they* impose themselves upon us, in the manner of uninvited dinner guests, refusing to leave even as we try to get rid of them. We can freely choose not spread a rumor, or sing a song, or tell a lame joke. With a modicum of self-control, we can even choose not to use an annoying buzzword, hum a catchy tune, or copy a mannerism. But we cannot freely chose *not* to believe something, if we have been exposed to appropriate evidence ([Adler, 2002](#_ENREF_1)). In the case of blatant and palpable truths, exposed right in front of our noses, the infection rate is nearly perfect.

What about beliefs that fail to correspond with reality, i.e. misbeliefs? Most such beliefs are quickly weeded out when we stumble upon evidence that disproves them, or fail to find any evidence that confirms them. Others misbeliefs, however, may be unlikely to encounter any destabilizing evidence. They may be difficult to falsify or too obscure to be open to epistemic scrutiny, there may be taboos or practical limitations preventing investigation, or they may be coupled with other beliefs that prevent their falsification ([Dennett, 2006](#_ENREF_23); [Talmont-Kaminski, 2013](#_ENREF_64)). David Haig agrees that the most likely candidates of selfish memes are to be found in complexes of belief-memes: “The place to look for sophisticated adaptation and selfishness will be in coherent ideologies, large ‘asexual’ meme complexes that are transmitted as a unit with high fidelity of transmission” ([Haig, 2007, p. 63](#_ENREF_31)).

Sterelny, on the hand, is more skeptical of the usefulness of the meme’s eye view for explaining the virulence of certain misbeliefs, even though he is sympathetic to other applications of memetics (in particular to physical artifacts). The main reason for his skepticism is that, for him, an explanation of the popularity of misbeliefs should fall directly out of a description of the workings of the human mind, without any need for the *meme’s eye view*:

The crucial problem is one of human psychology: explaining why we find occult-force explanations credible. Once we find out why humans find credible explanations of their environment in terms of occult forces, what else is there to explain? ([Sterelny, 2006, p. 159](#_ENREF_62)) “the socially-mediated flow of false belief does not turn on the nature of the beliefs themselves but on the details of human psychology. Once we understand the psychology of religious belief, there is no phenomenon that a meme theory and only a meme theory can explain.” ([Sterelny, 2006, p. 162](#_ENREF_62))

Sterelny is right that what makes for an attractive misbelief depends, among other things, on particular and contingent properties of the human mind. But we will now argue thatcertain systems of doxastic memesmay exploit our cognitive foibles in ingenious and unpredictable ways, which cannot be simply read off of human psychology.

By way of analogy, consider that adaptations in the living world also depend (obviously) on the details of the selective pressure in the environment. For instance, eyes will only evolve in a transparent medium, or they would be completely useless. However, that does not mean that studying this environment will provide a deep appreciation of the adaptive design of the camera eye. Just as stable features of a physical environment may lead to robust adaptive design in a living organism, the stable features of human psychology may give rise to robust adaptive solutions. Moreover, as we will see in our case study about witch persecutions (see further), systems of misbelief adapt not just to the make-up of the human mind, but also to local historical and cultural conditions. This makes Sterelny’s reduction of the question of the pervasiveness of certain misbelief to “details of human psychology” even more problematic.

In earlier publications, we have documented the many ways in which systems of misbeliefs are remarkably resilient in the face of counterevidence, and display features that ensure spurious confirmations ([Boudry, Blancke, & Pigliucci, 2015](#_ENREF_9); [Boudry & Braeckman, 2011](#_ENREF_10), [2012](#_ENREF_11)). Systems of misbelief seem to exhibit cultural design. They are equipped with immunizing strategies and defense mechanisms, protecting them from critical scrutiny and refutation, and they have features that make them self—perpetuating.

Importantly, this design need not be authored. It is more plausible that such design emerges after several successive modifications and combinations, as believers stumble upon sticky gambits and useful immunizations, and the most resilient and persistent beliefs tend to propagate in a population. By unwittingly selecting the misbeliefs that are resilient, infectious and conducive to further spreading, believers are setting a novel evolutionary dynamic in motion. Misbeliefs exploit both our cognitive psychology and the local cultural context to become more irresistible, and less vulnerable to destabilizing factors. They gradually become better and better at deceiving us, and to promote their own propagation. The misbeliefs that survive this selection tournament and become part of culture ([Sperber, 1990](#_ENREF_56)), are the ones that tend to be more intuitive and appealing, more resilient in the face of destabilizing evidence, better adapted to rapid dissemination, more conducive to spurious evidence, less open to epistemic scrutiny ([Dennett, 2006](#_ENREF_23); [Sperber, 2009](#_ENREF_59)), and more likely to motivate credibility-enhancing displays that infect others ([Henrich, 2009](#_ENREF_34)). The philosopher Stephen Law compared irrational belief systems with “intellectual black holes” in which “unwary passersby can find themselves … drawn in” ([Law, 2011, p. 10](#_ENREF_38)).

## Doxastic Catch-22

In contrast with other forms of cultural design, we cannot merely assume that such evolved misbeliefs serve the interests of their hosts, or even of society at large. Though misbeliefs can bestow benefits on their hosts, many others are probably deleterious, because they dispose to wrongheaded courses of action. Misbeliefs are hard to recruit and harness for the benefit of human individuals or groups because, by definition, doing so would undermine doxastic assent. Deliberately choosing and honing your own misbeliefs, in the interest of your health or happiness, runs into a performative paradox, which we call the Doxastic Catch-22. On the one hand, in order to reap the alleged benefits of a misbelief, you need to sincerely endorse it. But it impossible to voluntarily *decide* to believe something, after having weighed the costs and benefits (this was the major problem with Pascal’s wager). On the other hand, as long as you are under the spell of a misbelief, you have no way of figuring out whether it really serves your best interests. For all you know, it might be dangerous and destructive. If sincerely endorsed and acted upon, therefore, misbeliefs may easily lead to dangerous courses of action. Misbeliefs also tend to ramify through our belief system, with unpredictable downstream effects. Sooner or later, they might dangerously clash with reality.

In the evolutionary dynamic of misbeliefs which we have just described, human interests do not figure at all, except in the limiting case that, say, any misbelief which immediately motivates suicidal behavior would quickly be selected out of the meme pool. And even then, the very act of suicide may broadcast the meme to new receptive hosts. But short of discrete and unwitnessed suicide, there is a still lot of space for deleterious effects. In effect, it would take a conspiracy of benevolent and wise deceivers through the generations to ensure that misbeliefs *never* take a nasty turn.[[6]](#footnote-7) Naturally, such a benevolent deceiver would not himself accept those misbeliefs, but would carefully select them on behalf of his tribe or community, in their own best interests. He would be like the prophet Bokonon in Kurt Vonnegut’s *Cat’s Cradle*, confecting systems of bittersweet lies for his flock, without being able to reap their benefits themselves. Absent such implausible benign deceivers, there is only cultural evolution to put trust in. And what has evolution ever done for us? *Cui bono?*

As it happens, examples of such beliefs abound in the anthropological literature, but they are not always given due attention, in part due to the influence of functionalism. In his book *Sick Societies,* Robert Edgerton ([1992](#_ENREF_26)) has collected numerous examples of beliefs and attendant practices that gravely harm the interests of their hosts. Many of those relate to supernatural misbeliefs. In almost every known culture in the world, disease and other calamities are attributed to supernatural ghosts, demons, gods and witches. In some cultures, there is simply no other theoretical resource for explaining misfortune. In many societies, this belief in supernatural evil leads to time-consuming, wasteful, or outright dangerous behavior.

## Case Study: European Witch Persecutions

In this paper, we have no space to develop a case study of systems of misbelief in detail. One particularly promising test case for the meme’s eye view, however, is the phenomenon of witchcraft beliefs in early modern Europe. Belief in evil supernatural creatures abound all over the world, but early modern Europa witnessed the rise of a particularly virulent strain of ideas about witchcraft that motivated the persecution of tens of thousands of innocent victims. Witches, it was widely assumed, were ordinary Christians who had made a covenant with Satan, giving them the power to cast evil spells. A popular idea was that they convened at regular times during witches’ Sabbaths, usually by riding a broomstick or a goat, plotting evil plans to ravage communities. ([Goodare, 2016](#_ENREF_29); [Levack, 2006](#_ENREF_39))

Why did belief in witches, and the tendency to persecute them, spread far and wide, despitethe fact that there are no such creatures? If we want to understand the rationale of witchcraft beliefs, we cannot just ask the witch hunters themselves. It is useless to consult *them* about the function of their belief in witchcraft. Our question would baffle them, because the answer would be obvious from where they are standing: “Because there really are witches out there, of course, and we had better be aware of their evil ways!” And indeed, belief in the magical powers of diabolical forces made perfect sense in the intellectual world of 16th and 17th century Europe ([Clark, 1999](#_ENREF_14)).

Historians and social scientists with theoretical ambitions have been puzzled by the persistence of the European witch hunts. In the past several scholars had tried to devise functional explanations for the witch trials ([Ben-Yehuda, 1980](#_ENREF_7); [Erikson, 1966](#_ENREF_27); [Thomas, 2003](#_ENREF_65)). Perhaps, some argued, witchcraft persecutions foster social cohesion and protect the rules of civil conduct, by punishing anti-social transgressors. If you are afraid of being accused of witchcraft by your neighbors, after all, you might think twice before being rude or nasty towards them. Or perhaps witch hunts were useful ways of channeling aggressive impulses unto individuals at the periphery of the group, rather than allowing all that negative energy to rampage through more vital parts of the social fabric. After all, early modern Europe experienced several highly stressful social transitions. Or perhaps witch hunts provided some solace by offering explanations for misfortune that had befallen the community, and that would otherwise be unbearably mysterious.

However, there is no plausible evidence that witch persecutions performed such services on a systematic basis, either to individuals or society at large, as is now generally acknowledged by most historians ([Goodare, 2016](#_ENREF_29); [Hehl, 1987](#_ENREF_33); [Scarre & Callow, 2001](#_ENREF_55)). On the contrary, witch hunts often eroded group cohesion, increased mutual suspicion and exacerbated existing social tensions, creating fears about possible revenge from as yet undiscovered witches: “during witch-hunts people tended to draw apart rather than pull together” ([Stark, 2003, p. 217](#_ENREF_61)).

Another approach to the puzzle of witch hunts is to search for an interested party behind the scenes who might have profited from the persecutions. Perhaps it was a means for the ruling classes to dominate the common folk, or a ploy used by men to oppress women ([Barstow, 1994](#_ENREF_5); [Muchembled, 1987](#_ENREF_47)). “Too well designed” wrote the Marxist anthropologist about the witch-hunts system; the whole enterprise functioned so smoothly that there must have been an underlying motive apart from the stated goals of the witch-hunters ([Harris, 1989, p. 236](#_ENREF_32)). Someone somewhere in those communities must have decided, more or less consciously, to hunt down some witches, and must have benefited from the ensuing course of events, or else they would not have occurred in the first place. The *cui bono* question would eventually unveil the hidden culprit.

If witch hunts were secretly orchestrated, however, we would expect them to exhibit certain telltale patterns, for example targeting specific classes of individuals. But history shows otherwise ([Behringer, 2004](#_ENREF_6); [Levack, 2006](#_ENREF_39); [Scarre & Callow, 2001](#_ENREF_55)). During outbreaks of witch persecution, the chain of accusations could develop in rather haphazard and capricious ways.[[7]](#footnote-8) The division between persecutors and victims that these theories would lead us to expect, was far from obvious; initiative for the persecutions often came from ordinary people and woman, and males as well as people from the ruling classes also ended up on the stake. Large trials looked like genuine panics. They also rarely occurred at the same place twice. As a key witchcraft-expert explains: “Normally people are not interested in burning their neighbors, and governments do not wish to wipe out their tax-payers” ([Behringer, 2004, p. 149](#_ENREF_6)).

The meme’s eye view suggests a different approach, promising to shed new light on a conundrum that historians and social theorists have been wrestling with for decades. Doxastic memes about witchcraft were molded by cultural evolution, not to serve the interest of human actors, but to further their own reproduction and propagation. If you adopt the meme’s eye view, a number of features of witchcraft beliefs and persecutions suddenly start to make sense. For instance, the widespread notion of large Sabbaths, attended by witches in the region, implied that witches knew and were able to identify each other. Furthermore, belief in witches’ powers of flight lifted geographical restrictions on the attendance of the Sabbaths. Such beliefs facilitated long and prodigious chains of accusations.

Another portentous feature of witch hunts was the belief that it constituted as a “crimen exceptum”, a crime so very serious that the normal restrictions in torture could be disbanded. But people subjected to torture, in their desperation to end their ordeal, will confess to the most preposterous crimes. As one suspected witch confessed an ever longer list of accomplices under torture, officials in neighboring towns were warned of the infiltration of witches in their communities, leading to “spill-over” effects. As more and more people were convicted and burnt as witches, confessing to exactly the sorts of crimes that their torturers were expecting, the belief in witchcraft became more entrenched over time, while the magnitude of the persecutions became larger over the course of the 16th and early 17th centuries ([Levack, 2006](#_ENREF_39); [Monter, 2002](#_ENREF_46)).

From the meme’s eye view, the rationale of beliefs in large witches’ sabbaths was not the *reality* of such sabbaths, as the witch hunters believed, but the adaptive value such beliefs conferred on the fitness of witchcraft memes: they were more conducive to spreading the witch persecutions. Similarly, the memetic rationale of eliciting confessions under torture was not to discover real witches, but to trigger a chain reaction of accusations, ensuring that the witchcraft meme would live for another day. Specific beliefs about witches – their shape-shifting powers – were designed to make the belief system immune from falsification. The use of “spectral evidence” based on dreams and visions, for instance during the infamous Salem witch trials, provided a spurious source of confirmation. And of course, there is a glaring rationale to the meme that skeptical minds who doubted the existence of witchcraft, were themselves in cahoots with the devil ([Monter, 2002](#_ENREF_46)). Such memes, as Daniel Dennett writes, “disable the selective forces arrayed against them” ([Dennett, 1995, p. 349](#_ENREF_21)).

Now we can return to Millikan’s challenges. Millikan had argued that memetic evolution does not create novel purposes in the world, but merely satisfies the interests of their hosts in a more prodigious and efficient way than individual human ingenuity can muster. Cultural design may not be authored by human beings, but it still serves their interests (or the interests of some manipulator imposing memes on others). The example of witchcraft beliefs, however, offers a promising example of how a system of misbeliefs can evolve a functional rationale of its own which subverts the interests of its hosts. Beliefs in witchcraft have a purpose, but it is not represented anywhere in the minds of believers, and neither does it do them many favors. ([Dennett, 1995, pp. 78,164–165](#_ENREF_21)). In order to appreciate this design, we have to adopt the meme’s eye view, because the memes are the only suitable repositories of purpose.

# Discussion

There are several reasons why memes have not been taken seriously, and not all of them are bad reasons. The overblown claims of *panmemetics*, with their occasional outbreaks of Darwinian paranoia, have not increased the fitness of the *meme* meme in the academic world. The success of the meme meme in one environment has even undermined its fitness in another. Precisely because memes have become a staple of popular culture, with millions of Google search results, many academic researchers have treated them with suspicion and disdain. On top of that, there has been a problem of *linked loci*: because of its origins as an explicit analogue of the gene, the meme meme has been associated with atomist approaches to culture, and the notion of simple copying through imitation. Most of this criticism stems from an outdated conception of what a gene is, and a failure to appreciate the value of idealizations. In fact, cultural theorists critical of memetics have adopted units of culture of their own, which idealize and simplify human cultural in ways similar to the concept of a “meme”. But for the reasons mentioned above, they have resisted Dawkins’ nomenclature.

This would be mostly a semantic issue, if not for the fact that the most important element of the meme concept has been buried along with the term. This is the idea of *selfish* cultural elements, or mind parasites, which spread to benefit no one but themselves. In this paper, we have tried to unpack this claim with some care, taking issue with more permissive uses of the *meme’s eye view* in the literature, in particular with the grandiose claims of panmemetics. As signposts on this treacherous terrain, we have used challenges to memetics as voiced by its most thoughtful critics (Millikan, Pinker, Lewens, Sperber).

In the end, we argued that systems of misbelief may evolve into full-fledged mind parasites, developing novel purposes of their own, and subverting the interests of their vectors (not just of the latter’s genes). Fundamentally, this divergence of human and memetic purposes arises from the Doxastic Catch-22: it is conceptually impossible to steward your misbeliefs for your own benefit, or that of your community. If you start thinking about what your beliefs are ‘for’, or what benefits they might confer on you, they no longer have any sway over you. Because of this lack of reflective awareness, systems of misbelief may grow unruly and develop an evolutionary dynamic of their own. In the end, they are selected for reasons that remain opaque to their hosts. People will then be furthering the spread of misbeliefs that, unbeknownst to them, are adapted to ensure their own propagation, even at the detriment of their hosts. Doxastic parasites spread not because they serve human interest, but because they are more salient, attractive, better shielded from critical scrutiny and refutation, more conducive to spurious confirmation, and more likely to elicit credibility-enhancing displays that infect other agentso.

We think destructive phenomena like the European witch hunts should be Exhibit A of the meme’s eye view. They serve as a proof of concept for the general framework, showing how novel memetic purposes may arise through cultural evolution of misbeliefs, crosscutting the interests of their hosts. Traditional approaches to witch hunts have failed to find substantial functional benefits accruing from burning members of your own community, nor have they identified culprits who might have benefited from such a course of events (except in some cases). Our proposal is to treat the witchcraft beliefs themselves as selfish agents, furthering their own interests. By doing so, certain patterns of cultural evolution suddenly start to make sense in ways that were invisible to traditional approaches.

By taking a pragmatic approach to the selfish meme concept, and showing in some detail when the *meme’s eye view* may be fruitful or even indispensable, hopefully we can defuse some hostility to meme-talk, as well as temper the enthusiasm of advocates of panmemetics. To throw out Dawkins’ brainchild with the bathwater, as many critics of memetics have done, is to deprive oneself of an important conceptual tool to understand culture.

# References

Adler, J. (2002). *Belief's own ethics*. London: MIT Press.

Atran, S. (2002). *In gods we trust : the evolutionary landscape of religion*. New York: Oxford University Press.

Aunger, R. (2002). *The electric meme: A new theory of how we think*: Cambridge Univ Press.

Aunger, R. (Ed.) (2000). *Darwinizing Culture: The Status of Memetics as a Science*: OUP Oxford.

Barstow, A. L. (1994). *Witchcraze: A new history of the European witch hunts*: HarperCollins.

Behringer, W. (2004). *Witches and witch-hunts: a global history*: Polity.

Ben-Yehuda, N. (1980). The European witch craze of the 14th to 17th centuries: A sociologist's perspective. *American Journal of Sociology*, 1-31.

Blackmore, S. J. (2000). *The meme machine*. Oxford: Oxford University Press.

Boudry, M., Blancke, S., & Pigliucci, M. (2015). What Makes Weird Beliefs Thrive? The Epidemiology of Pseudoscience. *Philosophical Psychology, 28*(8), 1177-1198.

Boudry, M., & Braeckman, J. (2011). Immunizing strategies & epistemic defense mechanisms. *Philosophia, 39*(1), 145-161. doi:10.1007/s11406-010-9254-9

Boudry, M., & Braeckman, J. (2012). How Convenient! The Epistemic Rationale of Self-validating Belief Systems. *Philosophical Psychology, 25*(3), 341-364. doi:10.1080/09515089.2011.579420

Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford: Oxford university press.

Boyer, P. (1994). *The naturalness of religious ideas : A cognitive theory of religion*. Berkeley (Calif.): University of California Press.

Clark, S. (1999). *Thinking with demons: the idea of witchcraft in early modern Europe*: Oxford University Press on Demand.

Darwin, C. (1998 [1859]). *The origin of species*. Oxford: Oxford university press.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*: Oxford University Press, Incorporated.

Dawkins, R. (1993). Viruses of the mind. In B. Dahlbom (Ed.), *Dennett and his critics: Demystifying mind* (pp. 13-27). Oxford: Blackwell.

Delius, J. D. (1991). The nature of culture. In M. S. Dawkins, T. R. Halliday, & R. Dawkins (Eds.), *The Tinbergen Legacy* (pp. 75-99). Dordrecht: Springer Netherlands.

Dennett, D. C. (1990). Memes and the exploitation of imagination. *The Journal of Aesthetics and Art Criticism, 48*(2), 127-135.

Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon & Schuster.

Dennett, D. C. (2001). The evolution of culture. *The Monist, 84*(3), 305-324.

Dennett, D. C. (2006). *Breaking the spell: Religion as a natural phenomenon*. New York, N.Y.: Viking (Penguin).

Dennett, D. C., & McKay, R. (2006). A continuum of mindfulness. *Behavioral and Brain Sciences, 29*(04), 353-354.

Durham, W. H. (1991). *Coevolution: Genes, culture, and human diversity*: Stanford University Press.

Edgerton, R. B. (1992). *Sick societies: Challenging the myth of primitive harmony*. New York & Toronto: Free Press.

Erikson, K. T. (1966). *Wayward puritans: A study in the sociology of deviance*: Wiley New York.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. Oxford: Oxford university press.

Goodare, J. (2016). *The European Witch-Hunt*: Routledge.

Haig, D. (2006). Intrapersonal conflict. In M. Jones & A. C. Fabian (Eds.), *Conflict* (pp. 8-22). Cambridge: Cambridge University Press.

Haig, D. (2007). The gene meme. In M. Ridley (Ed.), *Richard Dawkins: How a Scientist Changed the Way We Think* (pp. 50-65).

Harris, M. (1989). *Cows, pigs, wars, & witches: the riddles of culture*: Vintage.

Hehl, U. v. (1987). Hexenprozesse und Geschichtswissenschaft. *Historisches Jahrbuch, 107*, 349-375.

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and human behavior, 30*(4), 244-260.

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*: Princeton University Press.

Ingold, T. (2007). The trouble with ‘evolutionary biology’. *Anthropology Today, 23*(2), 13-17.

Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge: MIT Press.

Law, S. (2011). *Believing bullshit: How not to get sucked into an intellectual black hole*. New York: Prometheus.

Levack, B. P. (2006). *The witch-hunt in early modern Europe*: Pearson Education.

Lewens, T. (2015). *Cultural Evolution: Conceptual Challenges*: OUP Oxford.

Lumsden, C. J., & Wilson, E. O. (1981). Genes, mind, and ideology. *The Sciences, 21*(9), 6-8.

Lynch, A. (2008). *Thought contagion: How belief spreads through society*: Basic Books.

McKay, R., & Whitehouse, H. (2015). Religion and morality. *Psychological Bulletin, 141*(2), 447-473. doi:10.1037/a0038455

Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences, 29*(04), 329-347.

Millikan, R. G. (2004). *Varieties of meaning: the 2002 Jean Nicod lectures*: The MIT Press.

Monter, W. (2002). *Witch trials in continental Europe 1560-1660* (B. Ankerloo, Stuart Clark Ed. Vol. Volum 4: The period of the witch trials).

Muchembled, R. (1987). Sorcières, Justice et Société aux 16 et 17 siècles.

Norenzayan, A. (2013). *Big gods: How religion transformed cooperation and conflict*: Princeton University Press.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2014). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences, 1*, 86.

Pagel, M. (2006). Darwinian cultural evolution rivals genetic evolution. *Behavioral and Brain Sciences, 29*(04), 360-360.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Pinker, S. (2012). The false allure of group selection. *Edge, Jun, 19*, 2012.

Plotkin, H. C. (2002). *The Imagined World Made Real: Towards a Natural Science of Culture*: Rutgers University Press.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago (Ill.): University of Chicago press.

Scarre, G., & Callow, J. (2001). *witchcraft and magic in sixteenth-and seventeenth-century Europe*: Palgrave Macmillan.

Sperber, D. (1990). The epidemiology of beliefs. In G. Gaskell & C. Fraser (Eds.), *The social psychological study of widespread beliefs* (pp. 25-44). Oxford: Clarendon Press.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Cambridge, Mass.: Blackwell.

Sperber, D. (2000). An objection to the memetic approach to culture. In R. Aunger (Ed.), *Darwinizing culture: the status of memetics as a science* (pp. 163-173). Oxford: Oxford University Press.

Sperber, D. (2009). Culturally transmitted misbeliefs. *Behavioral and Brain Sciences, 32*(6), 534-535.

Stanovich, K. E. (2005). *The Robot's Rebellion: Finding Meaning in the Age of Darwin*: University of Chicago Press.

Stark, R. (2003). *For the glory of God: How monotheism led to reformations, science, witch-hunts, and the end of slavery*: Princeton University Press.

Sterelny, K. (2006). Memes revisited. *The British Journal for the Philosophy of Science, 57*(1), 145-165.

Sterelny, K., & Kitcher, P. (1988). The return of the gene. *The Journal of Philosophy, 85*(7), 339-361.

Talmont-Kaminski, K. (2013). *Religion As Magical Ideology: How the Supernatural Reflects Rationality*: Isd.

Thomas, K. (2003). *Religion and the decline of magic: studies in popular beliefs in sixteenth and seventeenth-century England*: Penguin UK.

Williams, G. C. (1992). *Natural selection*: Oxford University Press.

Wilson, D. S. (2003). *Darwin's cathedral: Evolution, religion, and the nature of society*: University of Chicago Press.

1. Notice that this worry about loss of human agency and autonomy is not confined to memetics. Other theorists of cultural evolution, even those critical of memetics, have encountered the same criticism. For example, Ingold dismissed Mesoudi’s ([2006](#_ENREF_44)) framework in the following terms: “In this topsy-turvy world, it seems, human beings are but the means by which traits propagate themselves in an environment.” ([Ingold, 2007, p. 16](#_ENREF_36)) Lewens ([2015, pp. 26-28](#_ENREF_40)) offers a good discussion of this objection, concluding that “worries about passivity are not damaging to memetics”. Nor, of course, are they to other evolutionary approaches to culture. [↑](#footnote-ref-2)
2. Dawkins himself, although generally cautious about memes and modest about the ambitions of memetics, hinted at such a rebellion in the closing words of *The Selfish Gene*: “We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators.” ([Dawkins, 1976, p. 215](#_ENREF_16)). See also Dennett ([1995, pp. 364-367](#_ENREF_21)) for a discussion of this passage. [↑](#footnote-ref-3)
3. Note that the role of deliberate choice in artificial selection should not be oversold. The domestication of the wolf started as a form of what Darwin called “unconscious selection” ([Darwin, 1998 [1859]](#_ENREF_15)), with little or no conscious stewardship on the part of breeders. Only later on did conscious human intentions become the main selective pressure, and even then only in some lineages. [↑](#footnote-ref-4)
4. A similar objection was voiced, though less constructively, by Atran: “The distinction between good and bad memes really has no theoretical import. … the way the distinction is set up counts as a moral tale that is aggressively atheist and embarrassingly intellectualist.” ([Atran, 2002, p. 240](#_ENREF_2)). [↑](#footnote-ref-5)
5. Indeed, a convinced anti-semite may even turn the tables on us, arguing that the ideas of universal human rights and human tolerance are pernicious. [↑](#footnote-ref-6)
6. We are aware that some theorists have proposed *cultural group selection* as a mechanism for delivering functional benefits. Even if the theoretical problems with group selection can be overcome, however, and its precise causal mechanisms can be fleshed out ([Pinker, 2012](#_ENREF_52)), we think it is important to insist on the *cui bono* question: who are what benefits from group selection? For instance, Norenzayan and his colleagues have argued that belief systems involving Big Gods have “pro-social” benefits, because they allow human beings to solve the problem of large-scale cooperation among strangers ([Norenzayan, 2013](#_ENREF_48); [Norenzayan et al., 2014](#_ENREF_49)). But though the term of art “pro-social” sounds warm and vaguely uplifting, there is a dark side to “pro-sociality”, as it is intimately linked to violence and intolerance towards outsiders, and also towards potential defectors in the in-group ([McKay & Whitehouse, 2015](#_ENREF_43)). It may be the case that what goes under the name of “cultural group selection” can more profitably be understood as a limiting case of memetic selection, which would explain why religious belief systems often take a nasty turn. But this argument will be left for another occasion. [↑](#footnote-ref-7)
7. Of course, personal animosities could have played a role in specific instances, as could other incentives like greed or jealousy, but such factors do not come even close to explaining the scale of witch persecutions. Even if such personal incentives can be pointed to, this still leaves unexplained the specific cultural context that provided opportunities to settle personal scores in the first place: why did people believe in witches? [↑](#footnote-ref-8)