

Four Arguments that the Cognitive Psychology of Religion Undermines the Justification of Religious Belief

Michael J. Murray

Over the last decade a handful of cognitive models of religious belief have begun to coalesce in the literature. Attempts to offer “scientific explanations of religious belief” are nothing new, stretching back at least as far as David Hume, and perhaps as far back as Cicero. What is also not new is a belief that scientific explanations of religious belief serve in some way to undermine the justification for those beliefs.

Do these contemporary cognitive models of religion show us that religious beliefs are nothing but “a trick fobbed off on us by our genes?” Our first reaction to such a question should be: well if they do, it is not clear how. These models, if correct, show *not one thing more* than that we have certain mental tools (perhaps selected for, perhaps spandrels) which under certain conditions give rise to beliefs in the existence of entities which tend to rally religious commitments. But pointing that out does nothing to tell us about whether those beliefs are *justified* or not. After all, we have mental tools which, under certain conditions, give rise to belief in the existence of palm trees and electrons. We don’t regard those belief forming mechanisms as unreliable, nor (typically) the beliefs formed as unjustified. So what is it about religious beliefs, formed by HADD (hypersensitive agency detection device—see below), our penchant for minimally counter-intuitive entities, our attachment to full-access strategic agents, and so on, that makes us think that religious beliefs are epistemically unjustified? This paper considers four arguments for this claim.

Argument 1

In the most widely endorsed cognitive account of the origins of religious belief, those beliefs are spawned—or at least nurtured—into existence by a mental tool often described as a “hypersensitive agency detection device” (HADD). HADD is triggered by various environmental stimuli including apparently purposeful motion, configurations of matter, or physical processes with no apparent natural

cause. When triggered, HADD spawns belief in unseen agents that are taken to be the causes of the motions, configurations or processes. Evolutionary theorists hypothesize that such a mental tool would be adaptive since it would lead our ancestors to be especially sensitive to cues that might signal a predator. Furthermore, it would be adaptive for HADD to err on the side of excessive false positives since the cost of these would, for obvious reasons, be much greater than the cost of any false negatives. And indeed it does err in this way, thus explaining our tendency to immediately assume that bumps in the night are caused by *someone*.

In light of the fact that HADD is hyperactive in this way, it generates many false positives and is thus unreliable. Of course, beliefs that are outputs of a mechanism known to be unreliable are unjustified, at least unless those beliefs have some sort of independent evidence supporting their truth. Thus religious beliefs, spawned by HADD and not supported by independent evidence (and that covers the religious beliefs embraced by most folks), should be rejected as unreasonable or unjustified.

While this argument might be modified into a formidable objection as it stands, it fails. The problem with it is that it treats HADD as a belief-forming mechanism whose reliability can be assessed without reference to the contexts in which it is activated. This is wrongheaded, since belief forming mechanisms will routinely vary in reliability as the context changes. My visual system is generally reliable in helping me form beliefs about my physical environment, except in low light conditions, or when it comes to things that are very distant or small, etc. Similarly, HADD is quite reliable as a belief-forming mechanism in some conditions and perhaps not in others. In fact, in the ordinary course of things, it seems that when HADD triggers beliefs in agents, those beliefs *are* reliable. When you hear footsteps in the hall outside your room, or a knock on the door, or a whistled tune, you form beliefs in unseen agents, and you are right. However, it is also true that HADD has a tendency to misfire under some conditions. Can we distinguish conditions under which HADD is reliable from conditions under which it is not? Perhaps. We need not undertake the task of identifying all of those conditions. Instead we can conclude for now that religious beliefs spawned by HADD are unjustified only if HADD is likely to be unreliable in the contexts in which religious beliefs are generated. No scientist or philosopher has presented us with any such argument and, in fact, it is hard to see how such an argument could be formulated.

Argument 2:

Whatever cognitive tools are involved in spawning religious belief, one thing is clear: the beliefs spawned are obviously mutually incompatible. Any cognitive tools which give rise to mutually incompatible beliefs in this way are obviously unreliable, and any beliefs arising from them would then have to be taken to be unjustified (at least absent some sort of independent evidence of their truth). As a result, religious beliefs arising from these tools cannot be reasonably accepted absent independent justification.

One aspect of this argument is obviously correct: religious beliefs across times and cultures are largely mutually inconsistent. But for this argument to succeed it would have to presume something that is just as obviously false, namely, that the mental tools identified by cognitive scientists give rise to these inconsistent beliefs *all on their own*. When environmental conditions stimulate HADD, Theory of Mind, memorability and transmissibility via minimal counterintuitiveness, etc., the outputs of these tools are still highly non-specific. HADD tells me there is “an agent”; my beliefs about what sorts of fauna inhabit these parts lead me to conclude that the agent is a bear or a tiger or the bogeyman. If you conclude that it is a bear and I conclude that it is the bogeyman, this doesn’t show HADD to be unreliable, it shows that my mom was wrong to teach me that there is a bogeyman. Likewise, no one doubts that divergent cultural traditions play an enormous role in giving religious concepts their specific contours. If the mutually exclusive aspects of these beliefs creep in from cultural sources, this does nothing to undermine the reliability of these cognitive tools, it just shows that the cultural traditions are false.

What this shows is that for this second argument to succeed, one would need to show that the mutually incompatible aspects of these beliefs arise from the workings of the cognitive tools, and not the accretions or specifications that are introduced through cultural influences. It seems unlikely that this is the case.

Argument 3:

Cognitive psychological accounts of religion can account for the origin of religious belief in a way that makes no reference to and requires no causal connection with supernatural reality. However, properly justified belief requires that the target of the belief be causally connected to the belief itself in certain ways. Since these accounts show us that none of those ways are in fact in play in the origins of religious belief, beliefs so generated are unjustified.

This is a complex objection and it merits a more detailed response than I can give here. I have space for only the following two points. First, philosophers who are concerned about questions of the justification of belief fall into two basic categories: internalists and externalists. Internalists argue that whether or not a belief is justified depends entirely on facts “inside the head” (for example, facts about the extent to which the belief is properly supported by other beliefs one has). Externalists argue that justification depends on whether or not the belief has the right relation to facts about the external world. For some, the right relation consists only in *reliability* while for others the right relation will require that there be *certain causal connections* between the external world and the belief formed. I point this out only to make it clear that even if these psychological accounts do show that there is no direct causal connection between religious beliefs and their target, only some epistemological theories would take that to be relevant to the justification of those beliefs.

Should externalist epistemologists who argue for the importance of causal connections between the belief and its target take religious belief formed by HADD (etc.) to be similarly unjustified? No. Perhaps God set up our environment and the course of evolutionary history in such a way that we come to have cognitive tools that lead us to form beliefs in a supernatural reality (let’s call this the “supernaturalist explanation of religion”). If that is the way things work, then my beliefs would have a connection, albeit an *indirect* one, to the target of the belief, and a connection of that sort would not undermine the justification of the belief. We can see this by way of an analogy: Jones is a candidate in a local election. The voters in this district don’t pay much attention to this election and typically don’t know who the candidates are until they enter the voting booth. Wanting to become known to the voters beforehand, Jones programs an automatic dialing machine which calls each home in the district and delivers a short message introducing Jones and his campaign slogan. Because of this, all of the voters become aware of Jones and his campaign. If Smith were to stumble into Jones’ campaign headquarters and find the machine, could Smith conclude that the beliefs of the voters are unjustified because they were caused directly by the machine rather than directly by Jones? Of course not. The *direct* cause was a self-contained mechanism. But since Jones was the remote cause of the machine doing what it does, the beliefs are perfectly well justified. For all we know, God is likewise the indirect cause of the religious beliefs we have—beliefs that are directly caused by the cognitive tools psychologists have identified.

However, some scientist critics of religious belief are not satisfied with this response. For them, and this is the second point, something else is at stake in this argument, which they put like this:

In responding to Argument 3, the religious believer might push back by saying that “for all we know” God is the indirect cause of our religious beliefs, and in this way a proper causal connection between belief and target is preserved. That might tempt the religious believer into thinking that there is no better reason to affirm the “naturalist” explanation of religion over the “supernaturalist” explanation of religion. But that’s not right. We do have reason to favor the naturalist explanation and the reason is: it’s more parsimonious! Thus, while it is possible that the supernaturalist is right, it is more reasonable to assume that the naturalist is right.

This argument aims to press the superiority of the naturalist explanation on grounds of simplicity. But while simplicity is a scientific virtue, it is a virtue only when all other things are equal, something that is rarely the case. When we choose between competing hypotheses, we also need to take into account the other theoretical virtues of the competitors including the extent to which the hypotheses cohere with our over-arching worldview. To borrow an example, it is, all other things being equal, simpler to assume that there are no other minds but my own (and that the behavior of other apparently minded things is caused by purely mechanical processes) than it is to accept that there are many minds. But we don’t accept the “one mind” over the “many mind” hypothesis because such a hypothesis doesn’t cohere with many other things we are committed to. The Christian might be committed to the idea that all humans have minds because they are created in the image of God. The naturalist/physicalist might be committed to the idea that things that are physically alike are alike in other respects (including mental ones) and thus that all humans (like me) have minds. In either case, other considerations override considerations of parsimony in leading us to our view.

How does this affect the argument above? The answer is: the theist might say that the belief that God is the remote cause of supernatural beliefs fits in quite well with other reasons they have for thinking that religious belief is true (e.g., that without a God, many things just don’t make sense: the existence of objective morality, why there is something rather than nothing, why the universe is fine-tuned for life, and so on). For them, that explanation

would be more reasonable even if not simpler. Of course, for the naturalist, things might be different. However, all of this shows simply that psychological explanations of belief of the sort we have been considering do nothing on their own to undermine the justification of religious belief.

Argument 4:

While some belief forming mechanisms honed by natural selection can be supposed to be reliable, others cannot. The ones that can are those that produce beliefs “visible” to natural selection—beliefs about the physical world generated by the operation of the senses, for example. If my visual system produces beliefs that are largely incorrect, natural selection will catch up with me. Because of this, I can have some confidence that my visual systems are reliable. But supernatural beliefs generated by HADD and other cognitive tools are not subject to the winnowing power of natural selection in this way. If these cognitive tools were to mislead us systematically about the nature of supernatural reality, natural selection would do nothing to cure us of these illusions. As a result, we have no reason to think that their reliability would have been similarly honed. We ought thus to regard religious belief so spawned to be unreliable.

This argument makes two fundamental mistakes. First, it assumes that natural selection can indeed winnow reliable from unreliable belief forming mechanisms. Unfortunately, there is no reason to think this. The only way that natural selection can winnow these belief forming mechanisms is by winnowing the behaviors that they produce. Behaviors, in our case, arise from the interplay of beliefs *and desires*. Because of this, false beliefs can be as adaptive as true beliefs *as long as they are paired with affective systems that, together with the false beliefs, give rise to adaptive behaviors like feeding, fleeing, fighting, and reproducing*. Since false beliefs can be as adaptive as true ones, there is no reason to think that natural selection will select for reliable belief forming mechanisms and against unreliable ones.

The second mistake is that even if our imaginary scientist is right about the role evolutionary pressures play in giving us true beliefs about our environment, there is no reason to think that evolutionary pressures would lead us to false beliefs concerning religious reality. If the “supernaturalist explanation of religion” is correct, then our coming to believe that there is supernatural reality is something that leads us to truth belief because those beliefs are true. God set up the natural conditions so that, *pace* the objection, natural selection does select for reliable religious belief forming mechanisms.

