

ON MCCAULEY’S WHY RELIGION IS NATURAL AND SCIENCE IS NOT: SOME FURTHER OBSERVATIONS

by Gregory R. Peterson

Abstract. Robert McCauley’s Why Religion Is Natural and Science Is Not provides a summary interpretive statement of the standard model in cognitive science of religion, what I have previously called the HADD + ToM + Cultural Epidemiology model, along with a more general argument comparing religious cognition to scientific thinking and a novel framework for understanding both in terms of the concept of the maturationally natural. I here follow up on some observations made in a previous paper, developing them in light of McCauley’s own response to my previous arguments.

Keywords: Justin Barrett; cognitive science of religion; HADD; maturationally natural; Robert McCauley

Robert McCauley’s (2011) Why Religion Is Natural and Science Is Not provides a summary interpretive statement of what has become the standard model in cognitive science of religion, a model whose main features are shared by Justin Barrett (2004), Pascal Boyer (2001), Jason Slone (2004), and many others. This model gives prominent place to the theory that humans are innately prone to hyperactive agency detection and that the spread of religions can be explained on an infectious disease model. As McCauley argues, this is linked as well to the human capacity to conceptualize and simulate other minds, conventionally called “theory of mind” (ToM) in the psychological literature (Premack and Woodruff 1978; Carruthers and Smith 1996; Goldman 2006). McCauley’s version of this model is, as
his book’s title indicates, part of a broader thesis concerning the comparative naturalness of religious and scientific cognition. My analysis here is focused mainly on the theory of religion, and builds on a previous analysis to which McCauley has generously responded (McCauley 2013; Peterson 2013). McCauley’s _Why Religion Is Natural and Science Is Not_ (hereafter WRNSN) is highly interesting, probing, and provocative. But I have grown to believe that the theory of religion it argues on behalf of is significantly problematic if taken as anything approaching a complete theory, and if it is not intended to be a complete theory, it is difficult to see how it applies to those aspects of postindigenous religion that are of main interest. In this article, I focus on two lines of argument, the first concerning the language of naturalness employed by McCauley, the second concerning the model more generally.

**KEY CONCEPTS**

In the analysis that follows, three concepts in particular are important. The first of these is the maturationally natural (hereafter MN), a concept that McCauley devises to speak of features of human cognition that are (nearly) universal and which can at least be partly explained within an evolutionary framework. MN does heavy duty in WRNSN, since it is the basis for distinguishing the forms of religious and scientific cognition that interest McCauley. McCauley devises the concept of MN to address standard difficulties in traditional ways of speaking of the respective roles of nature and nurture in human development. In the past, and to a lesser extent in the present, it has been common to pit nature versus nurture, but it is a common understanding now that, for complex traits, what we conceive of as “nature” and “nurture” are so intertwined as to make distinguishing the two difficult. McCauley devises MN, it seems, to capitalize on relevant connotations of “nature” while also recognizing the pervasive role of “nurture” for any trait.

The second concept is that of hyperactive agency detection. As generally understood, such detection is enabled by a psychological module commonly referred to as a hyperactive agency detection device, or HADD (Boyer 2001; Barrett 2004). The HADD can be understood as a specific feature of the human capacity for ToM. On a standard account, ToM consists in the ability to think about the thoughts of others, and so to have a model of other agents that enables one to predict what other agents will do. The HADD concept implies that humans have a bias toward interpreting sensation in terms of agency and that such a bias would have an evolutionary advantage. As is commonly argued, I am better off thinking that the rustling of vegetation is caused by a tiger than by assuming it is caused by the wind, since the wind can’t eat me, but the tiger can. In the standard model of cognitive science of religion, HADD figures prominently as an
initial source for the generation of gods and spirits, and the implication is that the initial source of religion is mistaken detection: individuals think they see an agent when in fact they see none. In particular, they think they see “counterintuitive agents” referred to hereafter as CIAs. Such agents are understood to be “counterintuitive” in a very specific sense: they violate our own cognitive biases concerning what agents are like.

The third feature is that of cultural epidemiology (CE). McCauley adopts a disease model of religious transmission (Wimsatt 2010). Religions are understood to spread in the same way that diseases do: people are “infected” with religious ideas and such ideas are contagious because of the kinds of cognitive biases that we have. On a CE model, individuals are largely understood as passive hosts rather than active decision makers: it is the concepts themselves, memes if you like, that are doing all the work. In admittedly simplistic form, McCauley’s model of religion is one jointly of three concepts: HADD + ToM + CE. While HADD presupposes ToM, ToM does not by itself entail HADD, and neither concept (singly or jointly) entails CE. They are thus independent concepts that, when put together, might be understood to explain what religion is and how it propagates.

ARGUMENT 1: THE MATURATIONALLY NATURAL

My previous paper laid out two main arguments, the first concerning McCauley’s concept of the maturationally natural, and the second his
explanation of religion. In both cases, I laid these out in the form of a dilemma, and the dilemma I proposed to McCauley concerning the concept of the MN was this: either anything can be MN or the study of MN-linked processes is primarily the study of children and traits that develop in early childhood. In his book, McCauley emphasizes the early onset character of MN traits, and this raises important problems for understanding features of religion that are not cognizable or performable by children. But if MN traits are not limited to children, then there are many traits that would not be excluded but which don’t seem to be what McCauley has in mind. This is because a defining feature of MN traits is their (near) universality, and this is what distinguishes them from “practiced natural” traits such as the ability to write or riding a bicycle. But there are many beliefs such as “the sun rises in the east” and “only women give birth to children” which are universal but not obviously “natural” in any relevant sense.

In his response, McCauley seems to concede these points but does not fully address them, and he notes that he does discuss late onset maturational traits and has published on it elsewhere (McCauley and Henrich 2006). Left unaddressed is the question of what MN traits exclude in principle. To be “natural” in McCauley’s sense of the term requires that the trait be intuitive and fast as well as addressing fundamental problems that everyone faces. As already mentioned, they must also be universal. But this does not exclude the two belief examples already mentioned, and it is not clear that it would leave out such potentially universal traits of the future such as the capacity to surf the Internet or drive a car. These last two traits could be excluded if a further clause were added: that the traits in question must be cross-temporally universal, but this raises the risk that MN traits become too confined, and thus unable to explain features of modern religions that McCauley seems to think are explainable by his theory.

But I have also a more general concern, one not fully addressed in my previous analysis: does it really make sense to refer to these traits as natural at all? I did not raise this in my original analysis because this can be understood as a definitional question, and it is the sort of definitional questions upon which nothing of substance seems to ride. Philosophers use words in many different ways, and the important thing is that we use them clearly and avoid ambiguity. Arguments in this context about what counts as “natural” seem mainly rhetorical, and so (from the philosopher’s perspective) of lesser significance. Yet, I worry that McCauley’s thesis in places trades on rhetorical implications of the term natural that are contrary to his stated definition, and the number of responses that focus on McCauley’s use of the term “natural” demonstrate the conceptual confusion that can occur. The term “natural” in debates on human nature have a long association with claims about innateness and genetic dispositions, and McCauley has put himself in a position where he has to continually remind his audience that this is not what he means by natural at all.
This confusion could be resolved simply by replacing “natural” with “universal,” especially since it is the universality of such traits that are a key defining feature. To be a maturational trait, then, would just be a trait that is intuitive and fast (as opposed to slow and deliberate), and to be a universally maturational trait (UM) would just be such a trait that exists (nearly) universally. The contrast would then be with restricted maturational traits (RM) which do not exist universally, and these would be what McCauley has previously called “practiced natural” traits. One might then simply infer that if a trait is universal and maturational then it also solves fundamental problems of life, but if the inference is not obvious (and there is a fair chance that it is not!), then we could add a third term, perhaps “functional.” Thus we would have FUM and FRM traits, but also possibly UM and RM traits that are not F.

To redescribe McCauley’s concept of the MN in terms of the maturationally universal seems to change none of the content of his thesis while also capturing better (rhetorically) what he is trying to say. But it is also the case that, so redescribed, the thesis seems (again rhetorically) much less startling. After all, we already knew that religion is universal and science is not. We might not have known that religion is maturationally universal—this perhaps is new. But I’m not sure. It would seem that the startling features of McCauley’s thesis come not so much from claims about naturalness but the more specific claims about what religion is, how it originates, and how it is transmitted.

**Argument 2: HADD + ToM + CE**

The second set of arguments concerned McCauley’s account of religion specifically. The dilemma I posed was this: If McCauley’s goal is to provide a comprehensive account of religion, then the HADD + ToM + CE model is not up to the job, but if the goal is not to provide a comprehensive account, then the broader conclusions that McCauley wishes to make do not seem to follow, including the general claim that religion is natural. There are problems, for instance, with the HADD model of the origin of god beliefs. Experiments provide evidence that individuals preferentially remember CIAs, but they don’t show that people actually produce new CIAs to explain phenomena. Rather, if modest CIAs are provided, people simply remember them better than either normal agents or maximally counterintuitive agents. The experimental evidence about CIAs, then, seems at best to support an account of the transmission of some god beliefs, but not their origin. Since the God of the major monotheisms is not minimally counterintuitive, explanation of the transmission of the major monotheisms would be excluded. There are also problems with the idea that this preferential memory retention explains the transmission of religious ideas, the CE thesis. Here the problem is that there seem to be
well-known historical counterexamples, the spread of Islam and the origin
and spread of Mormonism as likely obvious ones. This is not to say that
the model that McCauley proposes, and which is widespread in the CSR
community, has no explanatory power at all; rather, I argue it has much less
explanatory power than McCauley seems to think, and it is perhaps most
problematic as an explanatory model for religion in literate societies.

McCauley agrees that his is not a comprehensive theory of religion,
and argues that such comprehensive accounts are beyond the scope of the
sciences (2013, 60). He also minimizes the significance of the HADD
account of religious origins. But he criticizes my description of his model
in terms of HADD + ToM + CE as overly simplistic, and he further
claims that I overly downplay the MN components present in Western
monotheisms. McCauley puts the point elegantly:

Why, for example, do people go on pilgrimages? Why do they both leave
things at and take things from pilgrimage sites (Nordin, 2009a, 2009b)?
Why do they pray to statues? Why are they more likely to pray to statues
that are nearby? Why do they cover statues’ mouths with scarves when air
pollution is bad? Why (in America) do they bury icons of Saint Joseph
in their yards when they want to sell their houses? Why do people open
Bibles randomly and drop their fingers to the page to find solutions to
personal problems? Pilgrimages, interactions with icons, and religious magic
are just three types of behaviors that appear to be grounded in theological
incorrectness. Ask the clergy. They know about this, because their job is
partly to police such behavior. (2013, 59)

Further, McCauley argues that the CE must be combined with a selec-
tionist model. So, to extend the epidemiology model, not only do religions
transmit horizontally the way that a disease transmits horizontally from
population to population, but just as diseases adapt to changing immune
systems, so too do religious systems. This, he argues, is a crucial feature of
the model, and he cites Jason Slone’s more recent work on sexual selection
(SS) theories of religion as one such additional selection mechanism.

The claim that I am downplaying the MN (or UM) characteristics of
religion and my charge that McCauley is overplaying them, especially with
respect to literate religions, is an empirical one. Let us take the example
of statuary that McCauley prominently features in his response. We can
ask the question: “How many Catholics seek to protect statues from air
pollution? How many Catholics bury statues of St. Joseph in their yard?
Further, what percentage of Catholics debates the fine points of limbo, or
the Trinity, or transubstantiation, and how do such debates affect their
behavior?” Let’s suppose that 60 percent of Catholics treat statues like
people or bury statues of St. Joseph. There is the further issue of what they
actually mean by such practices. Do they bury the statues because they
truly believe this has some effect, or do they do it simply out of tradition
(everyone does it!) or for some other reason that is not obviously MN/UM?
Let’s suppose that 60 percent of Catholics engage in these two practices, and that a full 50 percent of them do so for MN/UM reasons. Let us suppose further that 90 percent of Catholics engage in some MN/UM practices. Have we thereby explained Catholicism? Not remotely, for these practices may turn out to be rather peripheral to Catholicism, whether described by its practitioners or by scholars. Let us suppose that we can describe Catholicism as a set of beliefs and practices, perhaps the set “C.” We might use the Catholic Catechism or the online Catholic Encyclopedia as a source to make such a list. We could then ask what percentage of the beliefs and practices are satisfactorily explained by the model that McCauley supports. If the model was able to explain a full 90 percent of the beliefs and practices, that would be highly successful. Even 60 percent would be awfully good.

I have no idea what these percentages would be, but I have a guess. Meta-analyses of the literature of situational effects on moral behavior in psychology provide average correlations in the range of 20–40 percent, and I suspect that meta-analyses of relevant literature in CSR would show similar effects (Richard, Bond, and Stokes-Zoota 2003; Krueger and Funder 2004; Flanagan 2009). Such effects are real and important, but not determinative of behavior, and my hunch is that similar effect sizes will be found on average for those phenomena crucial to McCauley’s model, with either smaller or larger effect sizes in particular cases. Providing measures for effects of deliberative discourse and other non-MN/UM effects would be harder to assess, particularly since we would want to know longitudinal impacts and not simply short-term situational ones. Let’s suppose the impact of deliberation and deliberative discourse is a mere 40 percent. That would also be highly significant.

McCauley might object to using the Catholic Catechism to define Catholic beliefs and practices, as he frequently rejects official doctrine as a measure of religiosity. But I think we should find this perplexing. After all, many, many Catholics work through the Catechism at some level, and many, many Catholics would point to it as an authoritative source on outlining what Catholicism really is. It would be a strange thing to claim to explain Catholicism by explaining mainly those practices that Catholics do not find essential to Catholicism.

McCauley also charges that I portray his view simplistically, and I admit to some guilt here. Indeed, although I featured ToM as a central feature, it probably would have been better to highlight CIAs, since they play a more integral role in his framework. Initially, we might redescribe the theory as the HADD + CIA + CE model. But McCauley also gives considerable weight to the phenomenon of theological incorrectness (TI), a theory of ritual (TR) based significantly but not exclusively on his own work with Thomas Lawson, and a theory of myth (TM). If we add his endorsement
of Slone’s SS model and his modification of CE to include selectionist effects (CES) we have something like

\[ \text{HADD} + \text{CIA} + \text{TI} + \text{TR} + \text{TM} + \text{SS} + \text{CES} \]

It may still seem somewhat simplistic to some to frame McCauley’s theory this way, but I think it important to note the way that is in fact a combination of theoretical claims that, while having some affinity, are also truly discrete elements, and we should be careful in assessing the strength of the elements and their coherency. McCauley’s account of myth, for instance, is strikingly brief and short on empirical support. Like many practitioners of CSR, McCauley endorses both counterintuitive agency and TI, but I find the conjunction a little perplexing and well worth pondering. TI, a concept that has its source in the work of Justin Barrett and others and which has been developed in particular by Jason Slone (2004), is the idea that people deviate from official religious doctrine and that they often do so because of preset biases concerning religious belief. The idea of theological correctness fits well with the idea of the MN, since on this view we have “natural” religious tendencies that cause us to be more likely to adopt some religious beliefs over others.

It is important to note, however, that counterintuitive agency and TI are not identical concepts, and it is possible that their predictions run counter to one another. A famous experiment used in support of TI is that of Barrett and Keil (1996). The CIA model predicts that it is modestly counterintuitive agents gods that will persist in memory, and one might expect immodest CIA gods would not persist culturally and revert to modest ones. But the experiment by Barrett and Keil begins with documenting belief in an immodest CIA God, one who has more than simply one or two counterintuitive attributes, contrary to what the CIA model predicts. Subjects in the first two parts of the study read stories and then, on later recall, make inferences (and in some cases mistakes) that revert not to modestly counterintuitive agency, but normal agency. But if subjects make errors in the direction of normal agency, then one might expect that even retention of minimal CIAs should decay over time, contradicting other studies that show their persistence. Further, Barrett and Keil’s study show that this bias toward normal agency exists not for all CIAs, but it does exist for God beliefs even when people have time to deliberate and have the materials before them. So the bias (when it occurs) exists not only in intuitive processing but deliberate processing. Clearly, more careful thought and experimentation is needed in this area.  

I conclude with a final consideration of the respective roles of reasons and causes. Why do people believe what they believe? One possibility is that they may be caused to believe something, and these causes may act independently of (and perhaps contrary to) any reasoning processes. Much
of the language that McCauley and other CSR practitioners use strongly suggests a causal model. People perceive or think they perceive agents, and because of the design of their minds, they can’t help but do so. Some of these agents people come to believe are gods, and they are presumably caused by psychological mechanisms to do that as well. An inferential model, by contrast, would imply that some sort of reasoning process is taking place. Such a reasoning process could be implicit and unconscious, or explicit and conscious, or both. To what extent McCauley’s model (and those of other CSR practitioners) is causal/perceptual versus inferential/rational is of some import at the end of the day for determining how we assess people’s religious claims.

A causal model may seem to suggest that religious cognitions are false. If such causal processes are a byproduct of evolutionary endowments, this seems more likely, since without selection pressures it would seem that our religious cognitions would be accurate only by luck, and so we might expect only a 50 percent chance, a flip of the coin, that they are correct. If the processes are selected for, it may seem that the odds are greater, since we might expect that belief systems that are biologically selected for would be more likely to generate true beliefs. But the matter is complex, and the evolutionary psychology literature is replete with conjectures concerning how natural selection may preferentially select for false, self-deceitful beliefs, rather than true ones (e.g., Cronk 1994). Further, if God does exist, then it is possible and perhaps probable that God designed our cognitive faculties, and this would provide some ground for reliability within the context of a theological framework (Clark and Barrett 2011). A causal model doesn’t tell us whether God, or gods, or nirvana exists, but it does raise interesting questions on how we assess the reliability of religious beliefs.

Similar observations may be made concerning inferential models. We can ask whether the inferences made fit the norms of rationality (however defined). Inferential processes that fail to meet these norms would seem to produce true beliefs only by chance. But inferential processes may meet the norms of rationality and still produce many false beliefs. My cognitive faculties may be reliable, but only about 70 percent of the time, or they may be forced to operate in an environment other than the one they were designed for. If the inferential processes are complex, they would also be less predictable from a scientific standpoint. That people make inferences about whether God, or gods, or nirvana exist doesn’t tell us whether God, or gods, or nirvana exist, but we can then analyze the inferences people actually do use and assess them. We can ask whether they are justifiable for a given cultural environment, and we can ask whether and to what extent such justifiability extends across cultural environments. We might also model such inference procedures using the tools of decision theory. These are interesting questions, certainly for the philosopher, but also I expect for the psychologist.
This allows us to return to the previous question concerning the usage of “natural” with respect to religion. My own view is that there are both causal and inferential processes involved in the phenomenon of religion, and that any claim to full or nearly full explanation of religion must take into account both factors. Human beings have the cognitive capacity to be religious, and this distinguishes us from all other animals, chimpanzee rain dances notwithstanding. Many of us also likely have genetic or more broadly biological dispositions which contribute causally to favoring religious beliefs or certain kinds of religious beliefs. Some of us (perhaps many) may have dispositions that tilt the other way. So religion is partially, indeed very partially, natural in this genetic and biologically (innate) sense. But this is at best only part of the story. Human beings are also thinking creatures, and to model human beings without also taking into account this crucial facet is to simply model some completely different animal altogether. The model that McCauley proposes does not ignore this, to be sure, but I think it does underestimate it considerably, and much of the richness and complexity of religion, even among so-called “simple” folk, is lost.

A concluding statement: although my comments here are primarily critical, I want to affirm the importance of the cognitive science of religion as a field. Good and important work is being done. But I do worry about the limitations of current dominant paradigms, and I encourage ongoing efforts toward a broader consideration of approaches that better evaluate the complexity of religions and of human cognition, and so provide a better account of not what only what religions have in common, but also how they differ.

Notes


1. At least this is how ToM is conceived in terms of “theory-theory,” but a good case can be made in favor of “simulation theory” (cf. Goldman 2006; Keysers 2011). HADD as I understand it is conceptualized under the assumption of theory-theory, but the critique developed here does not depend on which account is correct.

2. McCauley (2013) has previously indicated his impatience with definitional questions, considering preoccupation with definitions as a “mopping up” operation. I have some sympathy with McCauley when refining definitions is used as a substitute for solving empirical matters. But I would argue that forming good definitions is important and integral to the progress of science, since doing so resolves ambiguity. The issue of defining and operationalizing “happiness” and its cognates in positive psychology is a good example.

3. “Practiced natural” raises its own rhetorical problems, as it seems to capture traits that are widely spread and automatic within a culture (but not across cultures) but are not practiced in an obvious sense. Thus, Americans share widely held convictions about the importance of political freedom. This belief is culturally specific, but not obviously practiced in the way that learning to ride a bike is. This would be even more obvious for such culturally widespread beliefs
that are not partially the result of conscious efforts of instruction, as would be the case with political freedom. Culturally specific standards of beauty come to mind.

4. Like others working in cognitive science of religion, it is important for McCauley that gods be understood as “minimally” counterintuitive, since experiments suggest that it is such minimally counterintuitive agents that are most easily remembered. A minimally counterintuitive agent is one which violates our expectations in only a few ways, for example, an “invisible man” or a woman who can fly. But the God of the major monotheisms has more than one or two counterintuitive properties. While the vocabulary of omniscience, omnipresence, omnipotence, omnibenevolence might be restricted to theologians and educated laity, recognition of the properties seems widespread. Other properties include invisibility, eternality and, more controversially for modern theologians, disembodiment and aseity.

5. McCauley may think that I have fallen into his hands at this point, since he does attempt to give explanations of both the Eucharist and baptism in WRNSN. But I find these unconvincing, and indicated this with respect to McCauley’s explanation of the Eucharist, which depends (in my view) on a misunderstanding of what Eucharist means for Catholics. McCauley’s response is that I am simply citing official doctrine. The argument seems circular, as we are being asked to believe McCauley’s explanation of the Eucharist because that is what his theory entails. But what is needed is some evidence that people actually believe what McCauley says they believe, and such evidence as far as I know is lacking.

6. McCauley does include consideration of later work by Cohen and Barrett (2008a, 2008b) on theological incorrectness in an Afro-Caribbean religious community, but this debate seems primarily a deliberative one, only indirectly (at best) relying on the fast, intuitive characteristics of maturational cognition. Barrett (2012, chapter 6) discusses his own research on this topic with populations in India and a similar unpublished study conducted more recently by Travis Chilcott and Raymond F. Paloutzian. The study by Chilcott and Paloutzian indicates age variation in response, with older Hindu populations being more anthropocentric.

REFERENCES