



The places of agency detection and predictive processing in the ontogenesis of religious belief; and “Who put the ‘H’ in the HADD?”

Andrew Ross Atkinson

To cite this article: Andrew Ross Atkinson (2023): The places of agency detection and predictive processing in the ontogenesis of religious belief; and “Who put the ‘H’ in the HADD?”, Religion, Brain & Behavior, DOI: [10.1080/2153599X.2023.2168731](https://doi.org/10.1080/2153599X.2023.2168731)

To link to this article: <https://doi.org/10.1080/2153599X.2023.2168731>



Published online: 30 Mar 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE



The places of agency detection and predictive processing in the ontogenesis of religious belief; and “Who put the ‘H’ in the HADD?”

Andrew Ross Atkinson  ^{a,b}

^aSociety & Cognition Unit, University of Białystok, Białystok, Poland; ^bDepartment of Global Development and Social Planning, University of Agder, Kristiansand, Norway

ABSTRACT

This paper addresses the origin of the hyperactive (or hypersensitive) agent (or agency) detection device (HADD) and its ‘places’ in the ontogenesis of religion. In 1. I engage in a preamble surrounding correspondence with Robert McCauley, Justin Barrett, and Tom Lawson on the origins of the HADD idea. In 2. I explain the roles HADD is purported to have had in the evolution of religion and a religious mindset from two potentially conflicting perspectives. I claim there is confusion where these two perspectives are not properly demarcated. In 3. I trace the evolution of the HADD concept. At 4. I outline the current condition of the HADD idea in light of several objections to it. I conclude that a) although history does not support the origins of the HADD idea as having any one single author, it is Barrett who spearheaded it and defended it, thusly galvanising his name to the concept historically; b) that as it is described in 2a it is compatible with predictive processing as described in 2b; and c) that the concept of HADD in its variant forms still excites philosophical and empirical questions, the answers to which are relevant to the science versus religion debate.

KEYWORDS

Hyperactive agent detection device; hypersensitive agency detection device; religion; cultural evolution; predictive processing; atheism

1. Introduction

During conversation with Tom Lawson some years ago after I had completed my doctoral studies, I expressed how I had very much appreciated the philosophical utility and implications of the idea of a hyperactive (or hypersensitive) agent (or agency) detection device (HADD). I had at one point been stuck with a mousetrapping problem. In effect, false positive agency detection seemed to lend itself to answering—if the atheist is correct about God’s existence—questions about how and why the concept of God appears (or rather begins to appear) in our ontology at all. I asked Lawson what I thought was a simple question (because I thought I already knew the answer and wanted to discuss it): “Who came up with the idea?” His answer was very clear: “It was Bob McCauley’s idea.” With Lawson, one can safely bet that what he says about the early days of the cognitive science of religion (CSR) is “gospel,” and so prior to writing this paper I had been citing him on this for the sake of historical accuracy any time I introduced HADD in my research. However, I eventually became frustrated at some reviewers’ comments about this. The complaints were always the same. “Is that correct?” HADD is most commonly attributed to Justin Barrett (Barrett, 2000). So to whom does the idea belong? I wanted to get to the bottom of it, so I asked McCauley directly, and rather to my surprise, he copied Barrett into the conversation along with Lawson. I now

had three of the main figures in the CSR engaged in a retelling of how one of the central ideas in the CSR came about, and their responses were each excellent—especially Justin Barrett’s, to whom I am particularly indebted here. I have been given their kind permission to relay some of the conversation here. The first reply to my question about who came up with the HADD idea came from McCauley:

This is the difference between informal conversation and formal publication. So far as I recall, Tom is correct that I first raised this idea in extended, intense conversation (at a 3-day conference that a big group of us were at in Finland), but Justin, so far as I know, was the first person to run with this idea in print. So, I suppose that I might be said to have invented it, perhaps, but Justin is definitely the person who knocked the idea into shape, developed the notion, and put it out for the public to ponder.

The second response came from Barrett, but, for now, I only quote it in part in order to clarify the introductory matter to hand. The rest of what he said forms the basis of the main inspiration for this paper. Barrett’s response to the question was as follows:

This is one of those things that I have no clear memory for and, having just read a review of how easy it is to acquire false memories, I am hesitant to trust any memory that I might conjure up. [...] All that said, I am confident that I was influenced by conversations with Bob and Tom because we frequently hashed out both theoretical and rhetorical issues concerning how to talk about this work we were doing.

The third response came from Lawson, and was typically Lawson in spirit:

I am quite aware of the distinction between vividness and accuracy of memory recall. I have a vivid memory of our discussions on the island Seili in Finland in 1999. Justin, Bob, Pascal, Harvey Whitehouse, Jesper Sørensen, Ilkka, and Stewart Guthrie were all participants in intense conversations. I have a vivid memory of one of the ideas that emerged and that first suggested by Bob was the idea of HADD. The discussions revolved around mechanisms that produced false positives.

Barrett then went on to say something much more important:

I do think that similar concepts – if *not* the term – were already in the cognitive and cognitive/developmental literature by the Finland conference in 1999, which Bob would have been aware of and brought into our discussions.

Indeed, the origins of HADD theory likely centered on shared focal themes implicated in the general scientific trends at the time. The upshot of the conversation however was twofold at least. On the one hand, my original question had been answered. On the other hand, they (Barrett in particular) gave me a lot more work to do because the story of who coined the HADD idea and how it came about is not so simple. Moreover, there appear to be multiple conceptions of it, some of which have been problematic, and perhaps still are. One is a shift from the use of “hyperactive” to “hypersensitive.” Another is the idea that HADD can “fire at nothing” and that the shift from “hyperactive” to “hypersensitive” represented either a move to alleviate that problem, or a suspicious shifting of the goalposts as if the latter term were a little more ontologically loaded in favor of positive theism.¹ Another is that it appears to rest on a modular approach to mind which is not popular with some scholars,² and another is that it is indeed a “device” at all. Another more serious problem which I will go on to address is what the HADD story was meant to be claiming exactly from its onset in the CSR’s conceptual repertoire. But does it matter who put the H in the HADD and what it stands for? Does it need to be modular? Does it even need to be “a device”? Should we be careful to specify what *sort* of agency we are attending to? Sensed presence? Causal animacy? Human-like agency? This paper attends both to the role HADD is purported to play in the evolution of religion, and to some of the theoretical precursors to HADD. The theory currently appears not to be in a good place according to some authors (Andersen, 2017; Barnes & Gibson, 2013; Tratner et al., 2020; Willard, 2017), and this is a shame because it has been one of the most useful ideas the CSR has deployed with important philosophical implications for stalwart problems in the philosophy of religion. I aim to defend HADD theory by arguing that perhaps many of its dissenters misunderstand its most crucial point, which I make in Section 2.1.

From the above correspondence, it is quite clear that although McCauley coined the term “HADD” it is Barrett who has, as McCauley says, “knocked the idea into shape.”³ In “knocking the idea into shape,” McCauley of course means that the concept has required defense, clarification, and revision from the outset—especially in the face of theories which would seek to refute it, supplant it, or render it obsolete. In what follows, I determine two approaches to understanding the role of HADD.

2. Two takes on the HADD

There are at least two approaches to understanding HADD, and in describing them, this section relays, initially, what the original purpose of HADD was supposed to be and its explanatory scope—or so I argue. I do so in order to isolate and immunize one of these approaches against objections to HADD coming from a new scholarly approach to the CSR discussed in Section 4. The same motivation underscores Section 3. Distinguishing between these two approaches is relatively easy to do, but it is also easy to slip between them; so one should take care to specify from the outset which one is attending to when discussing the role of agency detection in religion.

2.1. *Type I HADD: the primitive origins of supernatural belief*

It currently makes little sense to describe the role of HADD without describing in tandem the capacity for “mentalizing” about the intentional properties of agents. That is to say, if agency is detected by a mind that also has a faculty which makes inferences about its agent’s intentions and that agency detection indeed triggers it, then that extra faculty is also relevant. Two core cognitive competencies required for religious cognition are agency detection and theory of mind (ToM). The HADD is a primitive⁴ evolutionary mechanism said to have evolved as a means of minimizing false negatives in order to avoid more costly errors where agents are of real concern (e.g., such agents might threaten predation, attack, or theft of resources) (Barrett 2004, p. 31). ToM is a more sophisticated evolutionary adaptation particularly prominent in *Homo sapiens*. ToM plays a profoundly significant role in human affairs ranging from social relationships, to the ability to acquiesce in literary fictional pretence and entertain theistic ideas. Many in the cognitive science of religion hold that HADD and ToM played a crucial role in our evolutionary trajectory by generating key—even if rudimentary—religious concepts such as “supernatural intentionality” behind natural events. It is no surprise that a mind geared toward dealing with agency and intentionality might appropriate the world in those terms, thusly arriving at ideas that form the crux of most religious world views, such as a God with divine plan. This does not necessarily mean that HADD and ToM form the basis of all complex religious beliefs in the present to the occlusion of cultural processes (Barrett, 2011); rather, since we continually inherit these ancient cognitive mechanisms, evolved religious culture continues to be provided with fertile terrain, enjoying less fettered transmission than less intuitive sorts of information. This is an origin theory back down the long line of evolutionary time. Though the freeriding of religious culture upon these mechanisms may be said to account for religion’s maintenance in the present to an extent, the roles of HADD and ToM were geared more toward generative accounts of religion than the modes of its persistence. I call this *Type I HADD*. Expressed simply, from small ideas about intentional agency in our distant evolutionary past evolve bigger ideas, such as God; but this takes time. Type I HADD is not primarily aimed at explaining current religious proclivities; rather it explains the ontogenesis of a suite of cognitive traits that predates what most people refer to as “religion,” yet was crucial to the subsequent emergence of religion. Thusly, its explanatory scope is not strictly aimed at the ontogenesis of complex, present-day religious beliefs which always involve rich, culturally-elaborated content.

2.2. Type II HADD: a regular contributor to the maintenance of religious belief

A second, perhaps complementary take on HADD is that it is importantly implicated in current religious cognition and the ontogenesis of religious individuals, even if these processes are more directly and significantly influenced by cultural factors. I call this Type II HADD. Due to Type II HADD, a developing individual may well come to believe in very basic notions of supernatural agency during early stages of maturation, providing a fertile cognitive terrain that can, as the individual matures, be populated by more elaborated supernatural beliefs extant in the culture (Van Leeuwen and van Elk 2019). The Type II HADD view is generally assumed to play a role in the ontogenesis of a religious individual at a point in that individual's lifetime that precedes religious instruction, forming the basis of the given individual's cognitive religious repertoire and sense of the supernatural. Alternatively, Type II HADD could play a key role in generating powerful adult religious experiences—for instance, interpreting a lightning strike as a divine warning—that deepen believers' conviction in their existing religious beliefs, which are richly informed by cultural sources. Type II HADD's explanatory scope is not of the ontogenesis of religion itself over deep time, but of individual cases of religious maturation.

3. The convolution of the HADD idea, and Error Management Theory (EMT)

Now I want to turn to the evolution of the HADD idea itself. This is much different to describing the evolutionary rationale thought to account for the hyperactivity or hypersensitivity of the device as a means of minimizing false negatives. Here, the evolution of HADD refers to the steps leading up to its conception and development—this involved several twists and turns, hence, the “convolution” of the idea. For example, there have been a number of variants of HADD. In Barrett (2000) where the term first appeared, it is initially written “hyperactive agent detection device,” but in Barrett's subsequent work ‘hypersensitive’ and ‘agency’ ‘hyperactive’ and ‘agent’ respectively. Other authors have employed all possible permutations, based apparently on personal preference. On the matter of the “H,” Barrett says,

I recall that, at least in lectures, I first called this thing ADD with hyperactivity as a play on the psychological diagnosis. The H later migrated to the front because (1) it was easier to say and (2) the disorder had changed the name to ADHD.

By his 2004 book, “Barrett had realized that ‘agency (not agent) detection’ device was more appropriate,”⁵; in White (2021) both “hyperactive” and “hypersensitive” appear throughout the text.

There were indeed precursors to the HADD idea already in circulation before Barrett's coining of the term. Research into the perception of causal animacy has been in motion since Heider and Simmel (1944). As Lawson notes in various places, there were foundations in Noam Chomsky (Chomsky, 1957) and Jerry Fodor's revival of the modular theory of mind (Fodor, 1983). But Barrett himself does not appear to have been significantly influenced by these early predecessors: “I used this agency detection device as a substitute for and refinement of Guthrie's ‘anthropomorphism’ because I thought it fit better.”⁶ I also believe I was partly inspired by Simon Baron-Cohen's use of ‘devices’ and ‘mechanisms’ in his 1995.” (Baron-Cohen, 1995). Indeed, although different terminology is used therein, Guthrie (1993) cannot be ignored, and we absolutely cannot forget Charles Darwin's observations in *The Descent of Man* (1871) (Darwin 2009, pp. 65–69).

According to Barrett, Phillippe Rochat, also at Emory University with McCauley, had been working on developmental/experimental work on agent detection and agency attribution in the late 1990s (Rochat et al., 1997). Moreover, in the same issue of *Trends in Cognitive Sciences* where Barrett first used the term in print, there was an article on agency detection in children (Johnson, 2000), and a slightly later issue included a review article arguing that it was an encapsulated device or mechanism (Scholl & Tremoulet, 2000). According to Barrett:

This review paper along with some papers just earlier and later in the cognitive-developmental literature inspired a study I conducted with a student collaborator at Calvin College in early 2000 (Barrett & Johnson, 2003) concerning the agent detection device.

Also at the same time, another area of research was taking shape which was perhaps even more relevant to the HADD idea—EMT. Indeed, EMT sounds similar to HADD theory, but it does not involve modularity or devices, nor is it especially concerned with agency detection. EMT encompasses a broad range of strategies in evolutionary psychology for maximizing fitness by reducing errors of a certain kind. The theory holds that cognitive errors result from adaptive biases that exist in the present because they provided reproductive advantages for humans in the past. When it comes to agency detection, the evolutionary rationale for its hyperactivity or hypersensitivity is neatly captured by EMT. In Haselton and Buss (2000, p. 81) two types of error are hypothesized with one sort being less costly than the other such that the less costly error is likely to occur easily—in short, false positives. An example they use for the purpose of illustration is of a fire alarm being calibrated to sound out rather than fail, because the cost of the latter scenario is potentially fatal. This is directly parallel to the sorts of metaphors used by those seeking to explain the “H” in the HADD: e.g., “Better to mistake a stick for a snake, than mistake a snake for a stick.” Indeed, EMT appears to be implicit in the CSR literature (McKay & Dennett, 2009, p. 502). The beauty of appealing to EMT when talking about agency detection is that it avoids the need to talk of devices, or modularity of mind, and is not saddled with potential questions about whether or not the use of the terms “hyperactive” or “hypersensitive” come with any opaque philosophical undertones. The HADD is an error management mechanism for the detection of agency. Whether or not it is modular or distributed may well be irrelevant. The legwork that it does for the study of religion is to show how false positives in agency detection can occur quite naturally.

It is clear enough to see that there were multiple sources from which the notion of an agency detection device had begun to draw. Perhaps it is not so much a case of asking which individual scholar put the H in the HADD, but which groups of scholars were responsible for fleshing the idea out. The reason agency detection is hyperactive or hypersensitive is down to natural selection—in which case we have no reason to suppose anything more convoluted than the refinement of its sensitivity by trial and error.⁷ My particular philosophical worry is that when there is a general disarray in scientific literature about the meanings and use of terms, it will potentially deter new scholars from extracting what is useful about an idea in its proper form; seeing how it fits with other ideas, and moreover giving others a perceived licence to infer outlandish contrapositions (e.g., see Okasha (2010) for a similar concern surrounding research into altruism for a case in point). Perceived disagreements, and the subsequent adoption of misrepresentations and interpretations of the HADD idea have indeed led to some objections to it, and in some cases, attempts to jettison the concept entirely. I attend to some of these in what follows, with a view to focusing on predictive processing.

4. Objections to the HADD idea

When it comes to addressing problems with HADD, Barrett has to be the foremost authority. As mentioned, during the course of the email exchanges, Barrett presented me with a number of things to think about. I attend to those things in what follows, in some cases expanding on points already made above. There are places where I quote Barrett at length when what he has said clearly supports the arguments laid out in Sections 2 and 3. I believe doing so will be of interest to the reader in verifying certain positions, and I am also encouraged to discover that he and I are on the same page in many important respects, despite a difference in theistic outlook. Emphasis is Barrett’s own; citations are my own interjections.

(1) “It is sometimes claimed that there is no evidence for an agency detection device; it was merely postulated by evolutionary psychologists.” Barrett responded,

This is false. As I mentioned in our correspondence, that we have some sort of cognitive mechanism(s) for picking out agents from non-agents is an old idea with lots of experimental and developmental evidence. (Bloom & Veres, 1999; Heider & Simmel, 1944; Scholl & Tremoulet, 2000) The features of this mechanism may be disputed (e.g., how encapsulated it is, how much it generalizes to other sorts of things; does it only

pick out objects as agents or deal with traces or consequences of agency? *etc.*), but that there isn't [sic] one strikes me as pretty silly to deny at this point.

Indeed, to deny that we have an inbuilt capability to detect agency would be a peculiar claim. But, again, perhaps this matter is better seen through the lens of EMT and by taking the notions of its encapsulated modularity as a mere linguistic convention that isn't altogether necessary. Moreover, detecting cues in footprints in the snow or taking flight at the cracking of a sundried twig is quite different from the "sense" of agency. The "sense of agency" is perhaps as clear a thing as other less well-known senses such as proprioception (the sense of where you are in space), or equilibrioception (the sense of balance). I argue that a "sense of agency" is what the CSR hones in on when accounting for the origins of ideas about supernatural agency⁸ in a Type I HADD scenario, whereas detecting "signs of agency" is more akin to what some philosophers might draw on in support of physico-teleological arguments for creationism or intelligent design.

(2) "What does it mean for it to be 'hypersensitive'? We see slippage between 'hypersensitive' or 'hyperactive' being a claim about how the system manages error (when in doubt, which way does it lean?), versus metaphysical claims about whether it is a reliable system or not," writes Barrett. He continued:

It is for these reasons I've started dropping the H. [...] It seems most readers think I am claiming (following Guthrie, who does claim, I think) that HADD detects agents where there aren't any. I have specified that HADD sometimes detects agents/agency that is later overridden by the person and I remain agnostic as to whether it was right or wrong to override HADD.

A response to this is that it is perhaps a false dichotomy to deliberate on whether agency detection is reliable *or* not. It is both reliable *and* unreliable depending on context and what is meant. There is also something very different between sensing the living, and sensing the presence of the deceased during certain kinds of bereavement hallucination (McCauley & Graham, 2020, p. 64). Given that agency detection is *sometimes* unreliable, we should be skeptical about some of the beliefs which might result from it and, as Barrett says, perhaps override what we have no further evidence for. Relatedly, the sense of agency in the absence of adequate sensory input, and correcting for it, is perfectly accounted for by predictive processing, as argued below. While a sense of agency might in some cases begin with empirical cues that might indicate an agent being responsible for those cues, if there isn't an agent responsible, then agency detection may indeed be said to have fired at nothing. That is a matter of semantics. As for whether the change from "hyperactive" to "hypersensitive" represents any shift in conceptualization, some might claim that the latter indeed served to clarify whether or not HADD can fire without cue, arguing that it can't. I had previously thought it seemed ontologically loaded in favor of Barrett's own theistic beliefs. However, Barrett surprisingly said,

I don't believe I ever wrote about the change in what the HADD acronym stands for. At least as early as my 2004 book, I specify that the H stands for hypersensitive and the A stands for agency (p. 32). But in my 2000 TiCS article, the H is hyperactive and the A is agent. I don't know whether the change was with the book or before that.

Perhaps, therefore, it is a quirk of human reasoning that with a small change in an acronym, one assumes intentionality where there is none.

- (3) Some argue (e.g., Van Leeuwen and van Elk 2019) that there is little evidence that agency detection plays any role in supernatural agent beliefs. As has been argued, this problem disappears after making the distinction between Type I and Type II HADD. Barrett concurs, "This problem disappears if we take HADD to be significantly implicit in the evolutionary past of religion [generative]—not in every instance of relevant religious experience where it might be a response to prediction [maintenance]."

- (4) The fourth point Barrett raised related to confusion about the levels of analysis which my differentiation between Type I and Type II HADD endeavors to avoid. This he termed, “The difference between individual and group-level changes.” He said,

I accept some blame on this one. To illustrate the point that I meant to make (that something like HADD likely helps generate or at least encourage belief in supernatural agents following/adapting Guthrie’s thinking on this), rhetorically, it has been easiest to give examples of a person ‘detecting’ a ghost, spirit, *etc.*, and that leading to a personal belief that spreads. Some readers seem to have taken this to imply that for HADD to play a role in religious belief-formation, variation in HADD sensitivity or HADD experiences must correspond to differences in religious beliefs or at least religious belief intensity – particularly concerning God (of the Abrahamic sort). Of course, it doesn’t follow that HADD playing a role in *some* religious belief generation entails that it plays a role in generating *all* religious beliefs, such as the Abrahamic God. It also doesn’t follow that we each need our own HADD experiences to buoy our beliefs in supernatural agents. The claim that I was making (or at least trying to make) about HADD was that we’d expect that if HADD wasn’t tuned to accept pretty thin evidence of agency *and* happy enough for that agency to be non-human/unfamiliar/unspecified, then it would be much less likely for people to occasionally have HADD experiences that got attributed to superhumans; and then it would be much less likely for *groups* to be attracted to belief in such beings (and less reason to try to interact with them, *etc.*). And so, HADD becomes part of the fabric that leads to such beliefs on a group/cultural level, but not necessarily terribly important in the lives of individual believers or non-believers.

Apologies are of course due for the lengthy quote, but I wholeheartedly agree, and this quote verifies that Barrett has been focused largely on group-level ontogenesis over deep evolutionary time. The same is argued by Norenzayan (2013).

- (5) On the final point that Barrett raised, I wish to let him have his say:

I have seen that when people say there is “no evidence” for the role of HADD, what they have shown is that in some populations of individuals, measures of ‘agency detection’ sensitivity do not reliably co-vary with religious belief intensity. To which, I say, *so what?* That wasn’t the claim. I’d further add, that all of the studies I have seen of this sort have operationalized ‘agency detection’ as a species of ‘human-like agency detection’ such as finding faces, human-like movement in point displays, *etc.* This strikes closer to Guthrie’s ‘anthropomorphism’ than any claims about general HADD.

Barrett’s criticism applies to the following:

Several experimental studies have attempted to empirically investigate the HADD. van Elk and colleagues (2016) recently conducted a series of five experiments in order to investigate the effects of supernatural agent primes on agency detection. No effects of supernatural agent primes were found, with van Elk and colleagues also finding no general response bias towards agent detection as a function of heightened noise levels as would be predicted by the HADD model (van Elk, Rutjens, Pligt, & Harrel, 2016). The researchers, however, did find an increase in agency detection as a function of religiosity in three of the five experiments, which supports findings from previous studies. Riecki, Lindeman, Aleneff, Halme, and Nuortimo (2013) found that paranormal and religious believers produced more false positives than skeptics in a task where they were instructed to detect human-like faces in artifacts or scenery. Similarly, another study found that paranormal believers were more likely than skeptics to make false positives of agency under low and intermediate levels of noise in a biological motion perception task, although this difference disappeared at high noise levels (van Elk, 2013) (Andersen et al., 2019, p. 53).

Some apparent dissatisfaction with HADD, perhaps owing to misappropriations of it, have appeared to motivate alternative methodologies for accounting for religious agency detection via predictive processing. The role of predictive processing within the general CSR project has been in part a response to a perceived lack of experimental evidence in support of HADD. Proponents of predictive processing in the CSR would argue that perhaps what one means by false positive in agency detection in certain contexts is either a case of primitive agency detection or projection, and perhaps a subsequent particular failure to correct for it based on continuing information about the real world. On that view, one can still suppose an agency detection that isn’t perfect, while positing another mechanism that corrects for its errors further up the neuronal hierarchy (Andersen, 2017, p. 7). The study of predictive processing stands in contrast to the intuitive notion that the brain is merely a passive receptor of information from the external world. For some proposed examples of

the predictive mind at work, see Andersen (2017, p. 8). Another example which illustrates the predictive mind is the “rubber hand illusion” explored by Ehrsson et al. (2004). Another example that is particularly relevant here is the use of the mirror therapy (see Ramachandran and Blakeslee, 1998) to treat patients experiencing the peculiar sense of still “being with” the pain or paralysis of a severed limb. When pitched against HADD, some of its advocates claim:

while HADD theory claims that humans falsely detect agents by default and that this in turn gives rise to religious beliefs (Barrett, 2004; Guthrie, 1980, 1993) or strengthens them (Barrett & Lanman, 2008), the predictive processing model suggests that the reverse may also be true: religious teachings produce expectations in believers, and these then elicit false detection of agents, with one potential end result being the perceptual confirmation of cultural teachings and narratives (Andersen et al., 2019, p. 54).

I argue that the above-quoted position is characteristic of a Type II HADD misunderstanding of the Type I HADD point. That is, proponents of predictive processing appear to find HADD irrelevant in accounting for the generation of some supernatural beliefs and experiences⁹—but *when* is it irrelevant? We are dealing with opposite ends of a long evolutionary line. At one end attended to by advocates of predictive processing, it is likely correct to say that Type II HADD is not relevant, but that was never the point of HADD theory in principle. At the other end of the evolutionary line, predictive processing is equally irrelevant because the primers contained in the “religious teachings” and “cultural teachings and narratives” of interest to those authors had not yet evolved.

The two positions in fact likely intersect gradually over time. One way in which they might intersect is in the evolution of priors for agency detection (Maij & van Elk, 2017). On that view,

organisms are prepared to learn readily about phenomena that were relevant in an evolutionary past. For example, it is easier to condition people to fear animals, thunder, heights, and social events than to condition fear responses to modern threats such as cars or handguns (Maij & van Elk, 2017, p.92).

In the ontogenesis of religion then, the role of agency detection as an evolved prior would explain how religious ideas about supernatural agency proliferate.

Essentially, as Type I HADD becomes less significant, predictive processing becomes more significant. But as that time goes by, it neither renders the significance of Type I HADD null from an evolutionary perspective, nor does it render Type II HADD irrelevant for understanding contemporary religious belief. Proponents of predictive processing should perhaps argue that at the present end of the evolutionary timeline, cultural influences on an *individual's* religious ontogenesis are more significant than HADD for determining whether a given individual has religious concepts/beliefs/commitments. HADD theorists should perhaps concede that proponents of predictive processing in the CSR are absolutely right to have moved away from the explanatory bluster of a by-product account of religion. This is not intended to be a dismissal of predictive processing in favor of HADD; rather, it is intended to be a remedial effort to harmonize two theories that may be seen as talking past one another. Both agency detection and predictive processing approaches to the study of religion are valuable and perhaps equally indispensable. Andersen et al. do appear to concede this:

... while proponents of a HADD account claim that humans by default produce false detections of agents which in turn give rise to or strengthen pre-existing beliefs, a predictive processing account of agency detection suggests that the reverse direction of causality may also be true (Andersen et al., 2019, p. 60).

There is no significantly implicated theoretical difference, especially in the kind of “device” being implemented. That being said, it is perhaps the plight of certain scholars to maintain a strictly scientific approach to the role of agency detection in the CSR. These scholars approach the CSR without much talk of agency detection because they feel that Type I HADD is destined to remain an untestable hypothesis in the realm of the CSR’s philosophical foundations.¹⁰ Indeed, as Andersen et al. say, “The scientific value of these theories [...] depends on the development of proximate models that explain how underlying psychological mechanisms could produce such experiences in the first

place” (Andersen et al., 2019, p. 52). While this is merely a difference in methodology, it remains to be said that, upon trying to empirically test the role of agency detection in religious belief formation, one immediately slips into a Type II HADD way of thinking.

5. Conclusion

(a) The history of the HADD has been an interesting one, and it is clear that the term has been used to label something convolved from several areas of related research that agree that agency detection is an error management mechanism, whether modular or distributed. Where there are several antecedent pedigrees in the evolution of an idea, it is perhaps the case that some of those pedigrees have been bred together to form a new species of idea with some previous traits perhaps becoming vestigial in the process. With HADD, the ‘hyperactive’ or ‘hypersensitive’ H has been too costly to maintain due to the multifarious objections it suffered and it has been revised accordingly. Likewise, reliance on the modular approach to mind or the idea that agency detection has to be a ‘device’ in itself is no longer necessary for the idea to function in an explanatory role. Barrett himself has dropped the H altogether in favor of simply referring to “agency detection.”

(b) Even though it is an incomplete account of religious cognition, agency detection is nonetheless an important part the evolution of religious cognition. The predictive processing account of religious cognition holds that a certain kind of HADD is not relevant, the former being a better predictor of religiosity than the allegedly untestable HADD hypothesis. Proponents of predictive processing may hold that cultural forces are just as relevant, if not more so, for explaining core religious thoughts and behavior than any descriptions of the underlying cognitive mechanisms needed for them. However, that position is most forceful when Type II HADD is considered as it is in 2.2. If that position is valid, then the contraposition is equally valid. According to Type I HADD theory, the wealth of cultural forces attended to by predictive processing theorists are not relevant in our distant evolutionary past because their contents were not yet developed. HADD is only in some minor senses a maintenance theory of religion. Its most important place in the evolution of religion is as an origin theory.

Consider how a homeowner might come to believe that there is someone at the door when the doorbell is chimed. We have an agent detection device of a sort. Now let’s say that the occupant is informed that someone will call that afternoon, and so to listen for the doorbell. Perhaps that occupant insists on listening to noisy music with the effect that occasionally they rush to the door because they think they’ve heard the doorbell sounding (best to check). It’s not right to say that the doorbell is now irrelevant because the occupant can be made to believe there may be a caller by priming them beforehand. Thinking the doorbell has been heard even in error requires previous experience of its utility.

(c) Agency detection—in either its hyperactive, hypersensitive, modular, distributed, Type I or Type II forms—and predictive processing each make valuable contributions to broader issues stemming from the scientific study of religion, which can be addressed by those still interested in their philosophical implications. Does the origin of religious belief as perhaps resting on false positive agency ascription do damage to the claim that there is a God? Does the predictive processing model confirm the role of indoctrination in the making of religious individuals? The CSR has traditionally maintained that it is not concerned to answer that kind of question (White, 2021), yet such questions are still very much alive. It remains the case that although the traditional CSR position is to hedge on such matters, it could be a case of merely being polite about it (Boudry, 2021; Talmont-Kaminski & Atkinson, 2022).

It is highly likely that discussion of HADD in theological contexts will more frequently center on Type II understandings of it, especially where young earth creationism is at issue. Therefore, in addressing further potential discourse arising from Type II HADD misappropriations of the Type I HADD point, one should be clear that it is Type I that has been of greatest interest to those studying the evolution of religion and the role that agential cognition played in its ontogenesis.

Notes

1. By “positive” theism I mean the claim that “there is a God”; I allude to the distinction between positive and negative atheism, or explicit and implicit atheism (Martin, 1990).
2. For instance, “modules” (e.g. facial recognition) take input in a specific format (e.g. the facial information will work the right way up, but not upside down), but the way agency detection is often discussed appears to allow for a number of things ranging from sensed presence, to movement in point light displays, to facial recognition.
3. Barrett still maintains that he had been talking about agency detection during lectures well before the conference in Finland.
4. Agency detection might be supposed of many organisms, and is by no means a special attribute of *Homo sapiens*.
5. Barrett is referring to himself in the third person here.
6. Barrett elaborated that these early discussions of agency detection were meant to be a more cognitive specification of the sort of information processing that Stewart Guthrie had called anthropomorphism (Guthrie et al., 1980).
7. Barrett and Church go one step further in supposing that belief is warranted in a designer God as responsible for the various cognitive mechanisms which support religious belief, generally (Barrett & Church, 2014, p. 319).
8. I note here that some might argue that this approaches something like a *sensus divinitatis*, but it is at some remove indeed, for the latter (if it even exists) would surely come with a sense of reckoning and awe. The former is perhaps implied in the latter, but by no means does the former encapsulate the latter.
9. Some have argued that predictive processing can give rise to a cognitive bias that closely resembles the operations of the HADD (see, for example, Van Eyghen (2022). On such an account, it is no longer a modular, natural cognitive mechanism but it does reliably make humans hypersensitive in detecting agents.
10. This concern is indeed echoed by van Elk: “Without solid empirical evidence, the religion-as byproduct hypothesis remains nothing more than a “just-so-story”; i.e., a possible but not necessarily correct interpretation of our evolutionary past” (2016, p. 6). There is, however, good reason to move away from “just-so-story” concerns, toward abductive reasoning and inference to the best explanation.

Acknowledgements

I would like to thank Tom Lawson, Robert McCauley, and especially Justin Barrett for their permission to relay this material and their helpful comments. I would also like to thank Michiel van Elk for his input after the 2022 *Religion and the Mind* conference at Jagiellonian University.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research leading to these results has received funding from the Norwegian Financial Mechanism 2014–2021 (Project number 2019/34/H/H51/00654)

ORCID

Andrew Ross Atkinson  <http://orcid.org/0000-0002-5865-1514>

References

- Andersen, M. (2017). Predictive coding in agency detection. *Religion, Brain & Behavior*, 9(1), 65–84. <https://doi.org/10.1080/2153599X.2017.1387170>
- Andersen, M., Pfeiffer, T., Müller, S., & Schjoedt, U. (2019). Agency detection in predictive minds: A virtual reality study. *Religion, Brain & Behavior*, 9(1), 52–64. <https://doi.org/10.1080/2153599X.2017.1378709>
- Barnes, K., and Gibson, N. J. S. (2013). Supernatural agency: Individual difference predictors and situational correlates. *International Journal for the Psychology of Religion* 23 (1): 42–62. <https://doi.org/10.1080/10508619.2013.739066>.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.

- Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, 4(1), 29–34. [https://doi.org/10.1016/S1364-6613\(99\)01419-9](https://doi.org/10.1016/S1364-6613(99)01419-9)
- Barrett, J. L. (2004). *Why would anyone believe in God? (Cognitive science of religion)*. Walnut Creek, CA: AltaMira Press.
- Barrett, J. L. (2011). *Cognitive science, religion, and theology: From human minds to divine minds*. Templeton Press.
- Barrett, J. L., & Church, I. M. (2014). Should CSR give atheists epistemic assurance? On beer-goggles, BFFs, and skepticism regarding religious beliefs *The Monist*, 96(3), 311–324. <https://doi.org/10.5840/monist201396314>
- Barrett, J. L., & Johnson, A. H. (2003). Epidemiological and nativist accounts in the cognitive study of culture: A commentary on Pyysiäinen's innate fear of Bering's ghosts. *Journal of Cognition and Culture*, 3(3), 226–217. <https://doi.org/10.1163/156853703322336652>
- Barrett, J. L., & Jonathan, A. L. (2008). The science of religious beliefs. *Religion* 38(2): 109–24. <https://doi.org/10.1016/J.RELIGION.2008.01.007>
- Bloom, P., & Veres, C. (1999). The perceived intentionality of groups. *Cognition*, 71(1), B1. [https://doi.org/10.1016/S0010-0277\(99\)00014-1](https://doi.org/10.1016/S0010-0277(99)00014-1)
- Boudry, M. (2021). Why the genetic fallacy is not a fallacy. *Blog of the American Philosophical Association*. <https://maartenboudry.be/2021/06/why-the-genetic-fallacy-is-not-a-fallacy.html>.
- Chomsky, N. (1957). *Syntactic Structures* (1st ed.). Mouton de Gruyter.
- Darwin, Charles. (2009). *The Descent of Man, and Selection in Relation to Sex*. New York: Cambridge University Press.
- Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305(5685), 875–77. <https://doi.org/10.1126/science.1097011>
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Guthrie, S. (1993). *Faces in the clouds: A new theory of religion*. Oxford University Press.
- Guthrie, S. (1980). A cognitive theory of religion. *Current Anthropology*, 21(2), 181–203. <https://doi.org/10.1086/202429>
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91. <https://doi.org/10.1037/0022-3514.78.1.81>
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, 10(1), 47–66. https://doi.org/10.1207/s15327957pspr1001_3
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243. <https://doi.org/10.2307/1416950>
- Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences*, 4(1), 22–28. [https://doi.org/10.1016/S1364-6613\(99\)01414-X](https://doi.org/10.1016/S1364-6613(99)01414-X)
- Van Leeuwen, N., & van Elk, M. (2019). Seeking the supernatural: The interactive religious experience model. *Religion, Brain & Behavior*, 9(3), 221–251. <https://doi.org/10.1080/2153599X.2018.1453529>
- Maij, D. L. R., & van Elk, M. (2017). Evolved priors for agent detection. *Religion, Brain & Behavior*, 92–94. <https://doi.org/10.1080/2153599X.2017.1387591>
- Martin, M. L. (1990). *Atheism: A philosophical justification*. Temple University Press.
- McCauley, R., & Graham, G. (2020). *Hearing voices and other matters of the mind: What mental abnormalities can teach us about religions*. Oxford University Press.
- McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32(6), 493–510. <https://doi.org/10.1017/S0140525X09990975>
- Nola, R. (2012). Do naturalistic explanations of religious beliefs debunk religion? In G. Dawes & J. Maclaurin (Eds.), *A new science of religion* (pp. 167–193). Routledge. <https://doi.org/10.4324/9780203086131-16>
- Norenzayan, A. (2013). *Big Gods: How religion transformed cooperation and conflict*. Princeton University Press.
- Okasha, S. (2010). Altruism researchers must cooperate. *Nature* 467(7316): 653–655. <https://doi.org/10.1038/467653a>.
- Ramachandran, V.S., & Blakeslee, S. (1998). *Phantoms in the brain: Probing the mysteries of the human mind*. New York: William Morrow & Co.
- Riekk, T., Lindeman, M., Aleneff, M., Halme, A., & Nuortimo, A. (2013). Paranormal and religious believers are more prone to illusory face perception than skeptics and non-believers. *Applied Cognitive Psychology* 27(2), 150–55. <https://doi.org/10.1002/ACP.2874>.
- Rochat, P., Morgan, R., & Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cognitive Development*, 12(4), 537–561. [https://doi.org/10.1016/S0885-2014\(97\)90022-8](https://doi.org/10.1016/S0885-2014(97)90022-8)
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8), 299–309. [https://doi.org/10.1016/S1364-6613\(00\)01506-0](https://doi.org/10.1016/S1364-6613(00)01506-0)
- Talmon-Kaminski, K., & Atkinson, A. R. (2022). Science is disruptive, the science of religion particularly so. *Journal for the Cognitive Science of Religion*, 7(2), 213–225. <https://doi.org/10.1558/jcsr.20812>
- Tratner, A. E., Shackelford, T. K., Zeigler-Hill, V., Vonk, J. and McDonald, M. M. (2020). Fear the unseen: Supernatural belief and agency detection in virtual reality. *Religion, Brain & Behavior* 10(2): 118–131. <https://doi.org/10.1080/2153599X.2018.1526207>.

- Van Elk, M. (2013). Paranormal believers are more prone to illusory agency detection than skeptics. *Consciousness and Cognition* 22, 1041–1046.
- van Elk, M., Rutjens, B.T., Pligt, J. V. P., & Harreveld, F. V. (2016). Priming of supernatural agent concepts and agency detection. *Religion, Brain & Behavior* 6(1), 4–33. <https://doi.org/10.1080/2153599X.2014.933444>.
- Van Eyghen, H. (2022). Cognitive bias: Phylogenesis or ontogenesis? *Frontiers in Psychology* 13(July), 3881. <https://doi.org/10.3389/FPSYG.2022.892829/BIBTEX>.
- White, C. (2021). *An introduction to the cognitive science of religion: Connecting evolution, brain, cognition and culture*. Routledge Taylor & Francis.
- Willard, A. K. (2017). Agency detection is unnecessary in the explanation of religious belief. *Religion, Brain & Behavior* 9(1): 96–98. <https://doi.org/10.1080/2153599X.2017.1387593>