

*Belie the Belief? Prompts and Default States.*

Neil Levy

Macquarie University and University of Oxford

Department of Philosophy  
Macquarie University,  
NSW, 2109  
Australia

Uehiro Centre for Practical Ethics  
University of Oxford  
Oxford, OX1 1PT  
United Kingdom

neil.levy@philosophy.ox.ac.uk

**Abstract:** Sometimes agents sincerely profess to believe a claim and yet act inconsistently with it in some contexts. In this paper, I focus on mismatch cases in the domain of religion. I distinguish between two kinds of representations: prompts and default states. Prompts are representations that must be salient to agents in order for them to play their belief-appropriate roles, whereas default states play these roles automatically. The need for access characteristic of prompts is explained by their vehicles: prompts are realized in symbolic systems or even artifacts that make them inapt for automatic regulation of inference and behavior. I argue that some mismatch cases are explained by the fact that agents often report the contents of prompts when they report their beliefs, but behavior is controlled by prompts only when they are made salient to agents. I show that a number of otherwise puzzling findings in the cognitive science of religion, concerning belief intuitiveness, are illuminated by the distinction.

**Acknowledgements:** I am grateful to three reviewers for *Religion, Brain & Behavior* for extremely helpful comments that enabled me greatly to improve this paper. I am also grateful to an audience at the University of California, Santa Barbara, for their assistance in helping me to think through these issues. This work was supported by a generous grant from the Australian Research Council (DP180102384).

Beliefs are mental states with the function of guiding adaptive behavior. Apparently systematic failures of some beliefs to play this role are therefore highly puzzling. Such *mismatches* between what agents sincerely claim to believe and their behavior have in recent years given rise to a rich literature, in the cognitive science of religion, philosophy and beyond. This literature typically focuses on the extent to which the representations that explain agents' behavior in these cases occupy the functional roles of beliefs. Departures from these functional roles are cited as evidence that the agent does not fully believe, or believe in the ordinary way, what they assert. This approach has been used to assess the doxastic state of delusions (Bayne 2010); of the automatically activated 'implicit' attitudes that appear to explain racist and sexist behavior of those who profess egalitarian beliefs (Levy 2015; Mandelbaum 2016) and religious beliefs (Van Leeuwen 2014).

This standard approach has proven fruitful, but it makes an assumption that I will argue is not always true: that the functional role a representation plays is consistent across contexts. If this assumption is true, then apparent contextual variation in functional role reveals something about the properties that the representation possesses in *all* contexts, and a representation that fails to play the belief-appropriate role in one context cannot be a belief (unless the failure is somehow explained or excused). I will argue that consistency across contexts is not exhibited by all beliefs. Representations may play different functional roles across different contexts, in virtue of their *vehicles*, where the vehicle of a representation is the entity that carries the informational content, as opposed to the content itself (see Dennett 1969, 1991; Hurley 1998). Representations may be encoded in formats, codes or even media that entail that they play a certain role in some contexts and not others. These facts may explain some mismatch cases, without us being forced to conclude that the underlying representation is not a belief.

After introducing the kinds of cases that motivate the debate, I will turn to a distinction between two kinds of states that occupy the central functional role of beliefs: the guidance of flexible behavior. One kind of belief can be relied on to consistently play its distinctive role across contexts. The other kind should be expected to exhibit much greater context dependence, without that fact calling into question its status as a belief. I

will then suggest that different kinds of belief will tend to be realised by different kinds of vehicles, and that these facts about vehicles entail that there is a correlation between belief content, on the one hand, and vehicle (and therefore context dependence) on the other.

### *1. Mismatch cases and the current debate*

Mismatch cases come in a broad variety and have played a role in motivating several theoretical developments in psychology.<sup>1</sup> These cases range from the everyday to the pathological. At the everyday end of the spectrum, mismatches between what we take our beliefs to be and our behaviors have played an important role in motivating the postulation of implicit biases (Greenwald & Krieger 2006). At the pathological end of the spectrum, sufferers from monothematic delusions may profess bizarre beliefs while failing to act very consistently with these beliefs (Bayne & Pacherie 2005).

Within the cognitive science of religion more specifically, mismatch cases – mismatch between what religious believers are disposed to sincerely claim and their behavior – have recently attracted a great deal of attention. Such mismatches themselves come in a variety of forms. Some involve mismatches between normative commitments and behavior. Consider the *Sunday effect* (Malhotra 2008): Christians tend to behave consistently with their religious beliefs on Sundays more than other days of the week. For instance, the consumption of internet pornography is significantly lower in highly religious US states on Sundays, but not other days of the week (Edelman 2009). Analogous effects have been reported for Hindus and Muslims. Xygalatas (2013) found that Hindus who played an economic game in the context of a temple withdrew significantly less money from a pot that would otherwise benefit the entire group of players than those who played the game in a secular setting. Similarly, Duhaime (2015) found that Muslim shopkeepers in Morocco gave significantly more to charity within 20 minutes of hearing the call to prayer than at other times.

Other instances of mismatch in the religious domain do not appear to involve norms. Consider the phenomenon that has come to be called *theological incorrectness*. Theological incorrectness occurs when people who profess to believe the official tenets

of their religion appear to utilize inconsistent representations in interpreting religious stimuli; implicitly attributing limited knowledge or a limited capacity for attending to events to God, for example (Barrett 1999). Along related lines, there is evidence that afterlife beliefs are context dependent. Harris and Giménez (2005) and Astuti and Harris (2008) found that Spanish children and Vezo children and adults (respectively) were more likely to attribute continuing mental life to the recently dead when the person was described as dying in a context that featured primes for religion, such as an attending priest, rather than a more secular context.

These data seem to suggest that the representations that cause behavior in some, but not all, contexts in which their content is (apparently) relevant fall short of occupying the full range of functional roles we associate with beliefs. In the light of these kinds of data, we might conclude that they are not – quite – beliefs, or not beliefs of the same kind as more mundane beliefs, for that reason (Van Leeuwen 2014). We should think, that is, that despite their sincere utterances, the agents don't believe what they say they believe, at least not in the same way as they believe more mundane propositions. However, if the functional role of representations is sensitive to context in virtue of its vehicle, we should be wary of inferences from the role a representation plays to agents' attitudes. It may be the vehicle, and not the attitudes, that explains the context sensitivity. Such representations might count as genuine beliefs, despite this sensitivity.<sup>2</sup>

## *II. Two Kinds of Beliefs.*

Beliefs have the primary function of guiding behavior.<sup>3</sup> Organisms form representations of the external world because such representations give them a degree of flexibility unavailable if their behavior is under the control of reflexes or stereotypies. These representations come in a variety of forms, which enable greater flexibility as a function of at least two factors: the range of cues to which they are sensitive and the range of behaviors they are capable of driving. Following Sterelny (2003) we may distinguish *detection systems*, *robust tracking systems* and *decoupled representations*. Detection systems couple behavioral response to one or at most a narrow range of cues. For instance, a bacterium may follow a chemical gradient in water, or orient to the magnetic

field. Detection systems are extremely limited at both the input end (responding to only a single chemical, for instance) and the output end (driving only a single kind of behavior). Detection systems may be adaptive in informationally transparent environments, but to the extent to which environments are informationally noisy, there is selection pressure to develop robust tracking systems, which are responsive to multiple independent sources of information. Robust tracking systems enable greater flexibility of response by allowing the organism to respond across a wider variety of situations. But response is limited to a narrow range of behaviors. Decoupled representations are representations that are decoupled at both the input and output ends. They are formed or activated by a broad range of possible pieces of evidence (thus, the organism might form the belief “there is source of food” in response to olfactory cues, the behavior of other animals, colors, shape, and so on, rather than just a narrow range of cues) and they cause a range of different behaviors (perhaps pursuing the food in a variety of ways, perhaps strategic behaviors aimed at limiting the access of other animals to the food source, and so on).

Sterelny argues that decoupled representations are the best candidates for identification with the folk psychological notion of belief. There may be grounds for making some further distinctions, and thereby identifying a yet better candidate. A representation that was not merely decoupled at the input and output ends, but which integrated independent sources of information seems a better fit with our belief paradigm than one that does not integrate the independent sources. An organism integrates independent sources of information when it forms a representation in a way that calibrates the evidential value of different sources of information in the light of one another. Such an organism might, for instance, disregard an otherwise reliable cue for representation formation in the light of other cues that indicate that it is misleading.<sup>4</sup> Consider, for illustration, the representations that chimps form by monitoring the attention of other agents. Chimps are sensitive to various cues of attention: orientation of the body; visibility of the face; whether their eyes are occluded. But they have difficulty calibrating these cues in the light of one another. Instead, their performance in experiments indicates that some of these cues dominate others (so, for example, they prefer an informant who was oriented toward an event but had their eyes closed over one who was looking at the event but had their body twisted away from it). For a vivid illustration of how sensitivity

to a broad range of cues, combined with decoupling from specific response, is intuitively insufficient for genuine belief, consider dogs' response to 'invisible doors'. Dogs that have become familiar with glass doors treat them as insuperable barriers when the glass is removed, despite compelling visual (and olfactory) evidence that the barrier has gone.<sup>5</sup> An organism that is sensitive to multiple cues to a state of affairs but cannot integrate these cues with one another will have less behavioural flexibility and less fine-grained credences than one that can calibrate cues in light of one another. The best candidates for identification with belief, I therefore suggest, are decoupled representations that integrate multiple independent sources of evidence.

But such decoupled representations may themselves come in more than one distinct kind. Some of these representations are what I will call *default states*, and some of them are *prompts*. As the name suggests, default states govern behavior and inference automatically, without the need for special prompting or effortful recall. So long as fairly minimal conditions are met (the organism is awake and functioning within design parameters), behavior is responsive to the representation without the need for personal-level intervention. A *prompt*, in the sense in which I am using the term, does not automatically modulate behavior. Prompts must be *used* to guide behavior. Only when they are retrieved, primed or otherwise made salient will they play their characteristic role in cognition and behavioural regulation. When they are hooked up appropriately, they prompt behavior; hence the name.<sup>6</sup>

Most of our mundane beliefs about the world around us are default states. We don't need to be reminded or to make an effort to bear in mind that it is easier to exit a room via a door than through the walls, that pedestrians should take care to avoid fast-moving cars or that ink can stain. Some of our representations do not appear to be default states, however: they fail to govern inference and behavior across contexts. It is only when we are reminded of our religious commitments, for instance, that we are disposed to act consistently with them.

Prompts and default states have a great deal in common; enough to make it plausible to regard them as two species of the genus 'belief.' They are decoupled representations; there is no difference in principle in the range of behaviors they can drive. Again in

principle, they can be sensitive to, and integrative of, an equally broad range of independent sources of evidence. Prompts and default states also have in common other functional properties that have been proposed as distinctive of belief. Stich (1978) identifies *inferential promiscuity* as a central mark of belief; beliefs may serve as premises in inference by interacting with an indefinitely broad range of other representations. While prompts are defined by their failure to automatically modulate behavior, they may nevertheless satisfy Stich's criterion: *when* they are accessed, their content is poised to enter into inferential relations with other representations in just the same way as default states.

A further reason to regard prompts as a species of the genus 'belief' is that they satisfy the condition Rose, Buckwalter & Turri (2014) find to be the most powerful cue to belief ascription among ordinary people: frequent assertion. Asking people what they believe, under even minimally reflection conducive conditions, brings them to access their prompts and report their contents. It is, of course, the dissociation between such reports and behavior that gives rise to the puzzle of mismatch cases in the first place. Rose, Buckwalter and Turri find that for the folk a "persistent and robust profile of non-verbal behavior pales in comparison to the power of circumscribed but consistent verbal behavior" (692): if they are right, prompts, alongside default states, are "stereotypical beliefs".

There are, of course, other proposals for the functional roles definitive or characteristic of beliefs (see Zimmerman 2007; Jenson 2016 for proposals). While for reasons of space I won't go through these proposals here, prompts seem to satisfy these proposals just as well as default states, at least when they are accessed (and at least in principle). The fact that prompts must be accessed first does not seem to be good grounds for refusing to regard them as beliefs.<sup>7</sup>

### *III. The vehicles of beliefs.*

Default states may be evolutionarily ancient. There is evidence that at least some of the representations of other primates in our lineage are default states (though some of the same primates' representations may not be fully decoupled – they may be insensitive to

some cues for belief and may not be available to drive the full range of behaviors in the animals' repertoire; Sterelny 2003). Prompts may be evolutionarily more recent. It may be only in our species that representations which must be accessed to play the functional roles of belief exist. We have representational resources of kinds that other animals do not (or that they have only in a much more impoverished way), and these resources may be used to build prompts.

In particular, we have acquired representational systems; once we acquire such systems, we can acquire individual representations with novel contents; that is, contents that cannot be represented by those who lack such systems. The obvious example is natural language. In acquiring such a language, we acquire representational resources other animals have in, at most, a very impoverished form. Natural languages allow us to represent highly abstract concepts and form representations of events that occurred long before we existed and of places we have never been; such representations are difficult or impossible for non-linguistic beings.

We also have an acquired system of representing number and numerical relations. While the capacity to represent (very) small quantities precisely and large quantities approximately is *maturationally natural* (McCauley 2011) for us, the capacity to represent exact quantities without an acquired representational system is limited to very low numbers (Feigenson, Dehaene, & Spelke 2004). Mathematical representations that are trivially easy for us to form are inaccessible to those who have not acquired a number system.

That a representation is acquired using the resources of a representational system that was itself previously acquired does not entail that the representation is a prompt. Being a prompt is a synchronic property of a representation, not a historical property, and acquired representations may come to be automated (and thereby come to be default states). There is extensive evidence that representations that are historically dependent on acquired representational systems may come to be automated, at least partially. For example, there is plentiful evidence for semantic priming: agents process the meaning of words represented subliminally and the contents processed effect their subsequent behavior (McNamara 2005). Semantic priming is evidence of some degree of automation



of linguistic processing. The Stroop effect is also evidence for the automation of linguistic processing. When participants are required to name the color of the ink in which words are written, they are slower when the word is a color name and the color name is incongruent with the color of the ink ('red' written in blue ink, for instance), indicating automatic processing of the meaning of the word. There is also evidence of interference from numeral value for judgments of numeral size in the numerical Stroop task, indicating some degree of automatization with regard at least to low numbers (Pina et al. 2015).

However, this evidence supports only limited automation of the acquired representations. If, for instance, representations must be primed for them to play their characteristic roles – if mere relevance to the context isn't sufficient – these representations fail to be default states. There is, moreover, direct evidence that automation of acquired representations is often incomplete. For example, Dehaene et al. (1999) present neuroimaging and behavioural evidence that mathematical cognition depends on both visuo-spatial representations and linguistic representations. In their experiments, bilinguals performed better at exact addition problems when the problems were presented to them in the same language as a set of exact additions on which they had previously trained (whereas for approximate addition, performance was equivalent in both their languages). Bilinguals were also faster at generalizing their trained skill to new exact problems when they were presented in the training language. These data suggest at most incomplete automation of mathematical representations; were automation complete, accessing the linguistic representations would not give the person any extra benefit. Similarly, the evidence concerning semantic priming shows only partial automation. Priming fails for even two-word phrases, except when the phrase has been encountered often enough to have been overlearned as a unit. Negations cannot be primed, for example: the stimulus 'not good' primes 'good' (Deutsch, Gawronski & Strack 2006; Hasson & Glucksberg 2006).

The extent of automation of representations that are initially dependent on acquired systems varies across individuals and across domains. Dehaene et al. (1999) report that bilinguals showed no advantage of training language on approximate mathematical tasks – perhaps because these tasks piggyback on our innate approximate number sense – even

when the tasks involve cube roots and other relatively complex mathematical concepts. Further, the mathematical cognition of genuine experts appears to be much less vehicle dependent than that of non-experts (Almaric & Dehaene 2016). In other domains, however, even genuine experts show incomplete automation; Shtulman & Harrington (2016) suggests that experts in physics and in biology also have persisting folk physical and folk biological concepts that effect their cognition under cognitive load. To the extent to which such representations must be accessed in order to play their characteristic functional roles, they are prompts and not default states.

#### *IV. Explaining mismatch by prompts.*

We now have new resources to explain some cases of belief/behavior mismatch. Other things being equal, default states govern inference and behavior across contexts whenever their contents are relevant. Their contents are available automatically to play their belief-characteristic roles. But prompts aren't like that. For them to play their belief-characteristic roles, their contents must be accessed (effortfully or by being primed). If their contents are not accessed, then they will not control inference or behavior, even if their content is relevant.

This, I suggest, is what occurs in some of the parade cases cited in the literature as evidence that religious beliefs are different kinds of attitudes to factual beliefs. Consider the evidence that religious afterlife beliefs govern inference only when they are primed (Harris and Giménez 2005; Astuti and Harris 2008). If these afterlife beliefs are acquired and never fully automated, then they may be prompts and not default states. Such prompts would control behavior only when accessed. That is precisely what we see in these experiments: priming by experimental manipulation brings it about that these beliefs are salient and therefore available to play their belief-characteristic roles.

Of course, this explanation is persuasive only if (inter alia) there is good reason to think that religious representations are more likely to remain vehicle-dependent than other beliefs. Some people may find that claim hard to swallow: after all, religious belief is supposed to come easily to us. As McCauley (2011) says, religion is natural (whereas science is not). McCauley is explicit that by 'natural' he means *intuitive*: "intuition is the

principal manifestation of natural cognition in our mental lives” (2011: 14). But to say that something is intuitive is (again, according to McCauley) to say that it comes to mind transparently and rapidly (13). What is intuitive, it seems, is what can be expected to come to mind *automatically*. If religion is natural, it ought to govern cognition automatically, and therefore religious representations should be default states, not prompts.

Am I therefore committed to denying that religious beliefs are maturationally natural for us? It is worth remarking that the standard view in cognitive science is that religion is not a natural kind (Barrett & Trigg 2016). It is therefore quite possible – indeed, very likely – that some aspects of religion are maturationally natural and others are not. McCauley explicitly restricts his claim that religion is natural to what he calls popular religion. Doctrinal religion, religion elaborated by specialists, can depart very far from popular religion. It can be highly counterintuitive. The distinction between what is doctrinal (and potentially counterintuitive) and popular (and therefore intuitive) can do some work in explaining some mismatch cases. Consider the evidence concerning the phenomenon that has come to be called “theological incorrectness” (Barrett & Keil 1996). Religious believers who explicitly attribute theologically correct properties to God, like omniscience and omnipresence, appear to also have earlier emerging representations utilizing ordinary person concepts, which interfere with their inferences when they are required to respond under time pressure (Barlev, Mermelstein & German 2017; in press). The doctrinal properties of God that Christians fail to attribute in these experiments are counterintuitive, and therefore hard to think.<sup>8</sup> They may remain dependent on linguistic or other symbolic representations and are good candidates for being prompts. It is unsurprising that unless experimental participants are attending to their beliefs in these properties or had them primed, they default to more intuitive understandings of God.

While this response may explain why religious believers sometimes fail to deploy doctrinally correct concepts, however, and is therefore consistent with McCauley’s claim that popular religion is intuitive, many mismatch cases occur in the domain of popular religion. Dualist beliefs of the kind that fail to govern the inference of Spanish children (Harris and Giménez 2005) and Vezo children and adults (Astuti and Harris 2008) are widely held to be intuitive for ordinary people, and there is extensive evidence for folk

dualism (Bering 2002; Bloom 2004; Chudek et al. in press). The Vezo may not have a doctrinal religion at all; doctrinal religion is rare or absent without literacy (McCauley 2011). These findings seem to represent a challenge to the claim that afterlife beliefs are prompts, and not default states: if McCauley is correct in claiming that beliefs within the domain of folk religion are intuitive, we ought to expect them to come to mind without the need for prompting.

McCauley may be correct in suggesting that the precepts of popular religion are (often, at least) intuitive.<sup>9</sup> However, the precepts of popular religion may be intuitive in one, or both, of two quite different ways, and only one of these entails coming to mind transparently and rapidly. Representations that are intuitive in *that* sense – representations in accordance with which agents are disposed automatically to reason – possess what I will call *governance intuitiveness* (see Van Leeuwen 2014 for a related usage of ‘governance’). Default states possess governance intuitiveness, though other kinds of representations may also possess it. A representation that is governance intuitive plays its role (whatever that role might be) in cognition without the need for prior access and without drawing on scarce resources like attention. Prompts, however, lack governance intuitiveness, but they may be intuitive in a second sense: they may fit with our cognitive dispositions in a way that makes us apt to acquire them (call this *acquisitional intuitiveness*).<sup>10</sup>

McCauley conflates these two senses of ‘intuitiveness’ under the heading of “maturationally natural” cognition. The precepts of popular religion might be *maturationally* natural (with the emphasis on the former word) inasmuch as agents almost always acquire them in the course of normal development, even “in domains where they may have had little or no experience and no instruction” (2011: 5). On the other hand, they may be maturationally *natural* (with the emphasis on the latter word) inasmuch as they come to mind automatically and forcefully. While McCauley recognizes that intuitiveness of the latter kind may be the product of what he calls “practiced naturalness”, he appears to think that representations that are maturationally natural in the first kind are also maturationally natural in the second. This, I will suggest, is a mistake.

Many of the representations of popular religion are acquisitionally intuitive, but lack governance intuitiveness. In fact, the two major research paradigms in the cognitive science of religion – minimal counterintuitiveness (MCI) theory and hyperactive (or hypersensitive) agency detection device (HADD) theory – both entail that this combination of properties is likely. Both theories might be best understood as contributions to what Sperber (1996) calls cultural epidemiology. They are largely aimed at explaining the properties that make religious beliefs easily acquired, maintained; that is, acquisitionally intuitive. Insofar as they have implications for the content of the beliefs acquired, they suggest that these beliefs will lack governance intuitiveness.

MCI theory (e.g. Boyer 2000; Atran 2002) holds that the acquisitional intuitiveness of religious representations is due to their violating our intuitions – our cognitive grain – to a small extent. The talking serpent who tempts Eve in the Garden of Eden is MCI.<sup>11</sup> The serpent is counterintuitive inasmuch as he has the kinds of properties we associate with rational agents, like the power of speech. But he is only *minimally* counterintuitive in that the violation of our expectations is limited: the serpent is human-like in speech and reasoning, but doesn't possess any odd properties other than those that follow from possessing reason and the power of speech. There is extensive experimental evidence that MCI representations are at an advantage with regard to memorability: we are more apt to recall them than intuitive representations (even exotic ones) and we are more apt to recall them than more than minimally counterintuitive representations (see, for example, Boyer & Ramble 2001; Norenzayan et al. 2006).

We are more likely to acquire MCI representations. But the same properties that make them acquisitionally intuitive entail that they lack governance intuitiveness. They are, after all, counterintuitive (if only minimally) and that counterintuitiveness makes them inapt for inferences when they are not accessed. That is, those properties of the representation that make them counterintuitive must be accessed to play a role in cognition. In the absence of such access, people will tend to substitute simpler representations for their MCI concepts.

HADD theory, too, predicts acquisitional but not governance intuitiveness. It proposes that we have dedicated cognitive systems devoted to agency detection (Barrett 2004).

Agency detection is a critical task for animals like us: other agents represent crucial opportunities (for mating, trading, hunting) and threats for us. Crucial tasks like this, many believe, are often devolved to dedicated systems: modules (Barkow, Cosmides & Tooby 1992). Further, because agency detection is so vital, the costs of false negatives are much higher than the costs of false positives. Better to be needlessly vigilant than eaten by a predator. Our agency detection mechanism is therefore hypersensitive: it is sensitive to minimal cues of agency like apparent animacy or cues for eyes, disposing us to see agency at the (almost literal) drop of a hat. HADD theorists cite a range of evidence for the existence of the postulated 'device'. At least as far back as Heider & Simmel (1944), psychologists have accumulated evidence for how easily triggered our disposition to perceive agency is. Eyes are cued by stimuli as minimal as two dots on a piece of paper; triangles which interact contingently are attributed agency. According to HADD theorists, we are apt to generate, acquire or transmit religious representations in part because we are so sensitive to agency cues that we are disposed to see the hand of agents in the natural world.

Because we are so sensitive to cues for agency, supernatural beings possess acquisitional intuitiveness for us. But these same agents are often counterintuitive in other ways. They may be capable of violating physical laws, of combining the powers of different kinds of animals with speech and rationality, they may not be subject to death or decay, and so on. This combination of properties renders them acquisitionally intuitive, while predicting a very much lower degree of governance intuitiveness. A representation that has this set of properties is one we are prepared to recall, and (under the right cultural conditions) to accept. But it is not one that will govern our behavior automatically. It will require access for it to play the functional role of a belief. It is a prompt, not a default state.

As we saw above, there is evidence that afterlife beliefs do not govern behavior except when made salient *and* evidence that dualism is intuitive. This suggests that we have conflicting representations of death, with context (or chance) determining which representation governs inference and behavior. If one such representation is governance intuitive, we should expect it to be the default: that is, we should expect that representation to govern inference unless the conflicting representation is accessed. There is evidence that our biological representation of death occupies this default role.

Harris (2012) points to evidence that reasoning in accordance with afterlife belief increases with age. Young children utilize exclusively biological conceptions of death; as they age, the disposition to utilize religious afterlife beliefs increases. Outside contexts in which the religious afterlife belief is salient, however, we tend to default to our biological conception of death.

Despite lacking governance intuitiveness, dualism might be acquisitionally intuitive.<sup>12</sup> Perhaps it is in part *because* of the properties that entail it lacks governance intuitiveness that it is acquisitionally intuitive: these properties may render it MCI. We may be disposed to generate the concept of survival after the decay of the body, or to acquire it from those around us, or to transmit it to others. And when it is salient, it may govern inference. This hypothesis comports with the results reported by Chudek et al. (in press). Participants in their experiments (Canadian children and indigenous Fijian children and adults) watched an animation in which they were introduced to a pentagon identified by the experimenters as “Penny.” Penny, they are told, likes cake, and ‘she’ is seen apparently looking at, moving toward, and taking a bite out of a cake. They are subsequently shown an animation in which Penny is thwarted in her pursuit of cake by her inability to pass through a gap in a barrier. After apparently failing to pass through the gap, the pentagon stops moving, and a smaller triangle, till then motionless, becomes animate, looks toward the cake and passes through the gap. In the presence of strong cues for agency, most participants interpreted the narrative in a way that suggested that Penny had switched bodies, from the pentagon to the triangle; Chudek et al. suggest that this is evidence that a dualist perspective, according to which agents are distinct from bodies, is intuitive. Their data may indeed support their claim that dualist “intuitions emerge early and reliably in children everywhere”, but they do not support the claim that such representations possess governance intuitiveness. To be sure, the representations must be accessible enough to be made salient without special prompting: it was enough to introduce cues to agency and behavior that could not be explained except by the supposition of identity switches for dualism to be deployed spontaneously.<sup>13</sup> But this is consistent with thinking it would not be deployed at all in the absence of cues that render it salient.

It is less obvious how Bering (2002) is consistent with the framework proposed here. He presented participants with brief vignettes describing agents, at end of each of which the agent dies. Participants were then asked questions about the ongoing mental life of the dead person. As expected, ‘extinctivists’ – participants who claimed disbelief in an afterlife – were very much more likely to deny that the dead person had a continuing mental life. However, reaction time data suggested that they found it easier to give the extinctivist answer with regard to some mental states than others. The closer the conceptual link to the body, the quicker the extinctivist response. Thus, participants apparently effortlessly denied that the dead person was now hungry or tired, but took longer to deny that they still felt love or believed facts about their relatives.

Bering suggests that these data show that dualism is intuitive for all participants: at least with regard to the class of mental states that are not linked in folk psychology to the body, dualism is the default. It takes effort (and therefore time) to suppress the intuitive response in favour of the extinctivist response, despite the competition from the biological conception of death. This hypothesis is, however, consistent with the claim that dualism lacks governance intuitiveness. Dualism was highly salient in the context of the experiment; its influence on inference and behavior is therefore consistent with its lacking this kind of intuitiveness. Bering’s data is better understood as bearing on the relative strength of the representations when they are both salient rather than on whether either is governance intuitive. The dualist representation interferes with retrieval of the biological conception, slowing response.

What is the relationship between default states and prompts, on the one hand, and governance and acquisitional intuitiveness, on the other? Default states possess governance intuitiveness. They are poised to control inference and behavior across contexts without needing to be primed or effortfully accessed. We cannot, however, reduce the default state/prompt distinction to the governance/acquisitional intuitiveness distinction. Prompts may or may not be acquisitionally intuitive. Many or most of the representations that remain dependent on a symbolic system are unlikely to be intuitive in either sense of the word. They are acquired effortfully and govern behavior only when accessed.



We now have all the resources we need to explain a range of belief/behavior mismatch cases. When people are asked what they believe, they engage in effortful retrieval of the relevant information. Other things being equal, they are likely to report the content of their prompts, when they possess them. But these prompts will not govern their inference and behavior outside contexts in which they are primed or effortfully retrieved; rather, default states will play this role.

How far-reaching is this account? Can it explain all cases of belief/behavior mismatch? That seems highly unlikely. Not all instances of mismatch involve competition between prompts and default states. Some involve competition between two prompts; some competition between two default states (in which case relative strength of representation will play a large role in determining behavior). Nevertheless, the distinction between prompts and default states may be illuminating for some of these cases, too. Insofar as a representation is a prompt, we should expect it to be subject to contextual instability. What sort of representations are prompts? It is a sufficient (though perhaps not a necessary) condition of being a prompt that a representation is dependent on an acquired symbolic system. That, in turn, suggests that prompts are late developing representations; representations which we could acquire only once the symbolic system was in place. We should therefore expect that prompts are representations involving entities and relations of little adaptive significance in the environment of evolutionary adaptiveness, for it is these kinds of representations that had to wait on the development of symbolic systems.

If (as most researchers believe) religion is a byproduct of adaptations rather than an adaptation itself (Atran & Heinrich 2010), it is more likely to remain dependent on later developing symbolic systems and less easily automated.<sup>14</sup> At *very* least for those of us who do not achieve expertise within its domain, our representations are likely to remain prompts and therefore to control behavior only when accessed.<sup>15</sup>

### *Conclusion*

Sometimes, agents' behavior conflicts with what we should expect from them, given their sincerely asserted beliefs. Some of these conflicts, I suggest, are explained by the fact that

the representation they access in reporting their beliefs is a prompt. Since prompts control behavior solely when accessed, only when they are made salient or primed will these prompts play the functional role of belief. In other contexts, behavior will be guided by representations that may conflict with prompts. Since prompts are realized partly by acquired symbolic systems, we can expect such mismatches to arise in domains for which representations are likely to be dependent on such systems. There are a wide variety of such representations: in general, the greater the extent to which a domain is concerned with entities and relations of little direct adaptive significance in the environment of evolutionary adaptiveness, the more likely it is to be dependent on such systems. Many, perhaps the majority, of such dependent representations will never be fully automated, and will remain prompts.

It might be objected that prompts are not really beliefs. If prompts are not beliefs, then we have not explained belief/behavior mismatch: we have explained mismatch between states the agent *takes* to be beliefs and their behavior. Prompts have most of the properties that beliefs are taken to have. They are evidence responsive and inferentially promiscuous. But there are some grounds for denying that they are what we might call *full strength* beliefs; beliefs that possess all the properties characteristic of paradigm beliefs. We might think that full strength beliefs manifest the properties of evidence responsiveness and inferential promiscuity automatically; by definition prompts do not have that property. Perhaps such automatic regulation is part of what Schwitzgebel (2002) calls the dispositional profile of belief. Certainly, people seem surprised by belief/behavior mismatches, which is some evidence that we tend to think of beliefs as default states. But we shouldn't idealise beliefs (or agents): if we demand perfect conformity to the dispositional profile, we will end up with very few beliefs indeed. Should we think that automatic regulation is sufficiently central to the dispositional profile to conclude that states that are otherwise belief-like should be relegated to second class status, or expelled altogether from the category? If we think that automatic governance is required for a state to count as a genuine belief, but I am right in claiming that many mathematical representations (surely paradigms of evidence sensitivity and inferential promiscuity) don't have this property, perhaps that's a better reason to alter our sense of what it takes to be a belief, rather than to draw a line dividing prompts from default states.<sup>16</sup>

Terminology matters. It matters how we classify things, because taxonomy has knock on effects on how we think. It is, nevertheless, important to see that it is terminology that is at stake here: the distinction between prompts and default states is a genuine one, and it is one we can characterise in ways that avoid the question whether either deserves to be called a belief.

## Notes

<sup>1</sup> Elsewhere, I have advanced alternative explanations of mismatch cases (Levy 2017a; Levy in press(a); in press (b)). The heterogeneity of such cases suggests a correlative heterogeneity of explanatory accounts, so I am not (very) embarrassed to return to them here.

<sup>2</sup> Vindicating the status of these representations as beliefs is far from easy. The extent to which it is appropriate to identify a mental state with a folk psychological concept depends, inter alia, on our theory of reference as well as our analysis of folk psychological commitments, and I have nothing very illuminating to say in either regard. At very least, however, the considerations I will advance should suffice to forestall the claim that religious believers take a different attitude to the representations I consider.

<sup>3</sup> Arguably, guidance of behavior is not the sole function of beliefs. Sterelny (2015; in press) and Funkhouser (2017) have argued that beliefs may also play a *signalling* role. Perhaps beliefs can play other roles as well. However, these functions are likely parasitic on behavior guidance, which emerges (in some form) very much earlier in evolution. Agents who signal that they have a belief signal that they are disposed to act in certain ways.

<sup>4</sup> In fact, breadth of behavioural response cannot dissociate very far from a capacity to integrate independent sources of information, because behavioural control is dependent on the capacity to identify different dimensions of variation and assess them as relevant (see Levy 2017b).

<sup>5</sup> There are many videos of the phenomenon available online; see <https://www.youtube.com/watch?v=dZVF6hHUuBU&t=2s>.

<sup>6</sup> The default state/prompt distinction cuts across the personal/subpersonal distinction. Default states do not need to be recalled in order to govern cognition; hence subpersonal states can play the role of default states. However, some personal-level beliefs – those endorsed by, or otherwise appropriately attributed to the person – function as default states. Prompts may tend to be personal level states, that require conscious endorsement to play their characteristic role. However, priming can serve to activate prompts, and there may be subpersonal representations that can be primed. I thank a reviewer for this journal for prompting me to clarify my thinking on this point.

<sup>7</sup> Zimmerman (2018) argues that beliefs should be identified with those states that guide our attentive behavior, and that it is only when we attend to what we are doing that such states guide our behavior *qua* beliefs. While Zimmerman does not require that we attend to our beliefs for them to play their distinctive role, if we accept his proposal we will have even less reason to deny that prompts are beliefs, since prompts are the states that tend to guide attentive behavior.

<sup>8</sup> Recalling Feynman's quip that no one really understands quantum mechanics, McCauley suggests that the same "might well be said about all of the most famous doctrinal conundrums of Christianity (or any other doctrinal religion), such as those proffered for resolving the Christological and Trinitarian controversies of the early Church" (McCauley 2011: 153).

<sup>9</sup> While the view that some kind of religious thought emerges spontaneously, either as an adaptation or a byproduct of adaptations, is probably the most common view in the cognitive science of religion, there are dissenters who hold that a substantial amount of cultural scaffolding is required for the emergence of anything like a religion (see Gervais et al. 2011; Sterelny, in press, for this view). On no view are *all* the elements of religion acquisitionally intuitive. I am grateful to a reviewer for this journal for pressing me to clarify this issue.

<sup>10</sup> There is also a third kind of intuitiveness: a representation may strike us as obviously true only when we attend to it. Mathematical truths have this property, as does the Cartesian *cogito*.

<sup>11</sup> Purzycki & Willard (2016) argue persuasively that MCI theorists often conflate what they call *shallow* and *deep* inference. Roughly, shallow inference is inference over culturally acquired concepts; deep inference is over maturationally natural concepts. The serpent in the Garden of Eden may violate merely shallow inferences. Purzycki & Willard claim that actual religious beliefs are largely counterschematic (violating shallow inference) and not genuinely counterintuitive. The conflation of shallow and deep inferences does not seem especially problematic for the purposes of my project: I am concerned with the properties that make representations inapt for automatic control of inference, and violations of shallow as much as deep inference may make a representation inapt for such control. See Russell & Gobet (2013) for a defence of a non-nativist account of counterintuitiveness, which understands it as requiring conflict with exemplars which may be acquired (I do not, however, accept Russell & Gobet's claim that counterintuitiveness disappears as expertise develops; as we saw above, while expertise may suffice to cause some representations to possess governance intuitiveness it previously lacked, in some domains such intuitiveness is never acquired).

<sup>12</sup> The claim that dualism is acquisitionally intuitive does not commit me to the claim that it is natural, in the sense in which religion is supposed to be natural by Boyer (2008) and Barrett (2004): a cognitive default that will reliably emerge across almost all environments. It might be acquisitionally intuitive in that it is easily acquired given the right cultural scaffolding.

<sup>13</sup> It might be appropriate to identify another dimension of intuitiveness: the ease with which a representation is rendered salient. A representation that possesses *full* governance intuitiveness need not be primed at all. But some representations that fall short of this standard may be very easily primed, such that in most contexts in which their contexts are relevant they govern inference and behavior.

<sup>14</sup> This is even more likely to be the case if religion is the product of cultural evolution, to the extent to which cultural evolution is dependent on symbolic systems. Of course, even if religion is a genetic adaptation, it would not follow that religious representations were always or even often default states: the precise content of religious representations is culturally specified and therefore may be dependent on cultural resources. See Powell & Clarke (2012) for an exploration of the prospects of an adaptationist account of the origins, or maintenance, of religion.

<sup>15</sup> Barlev, Mermelstein & German (in press) report evidence that theologically incorrect representations compete for retrieval and control of response even among older adults

with extensive religious experience, suggesting that their religious representations remain prompts. Their subjects were lay members of Christian congregations. Replication with priests, theologians and other religious experts is required to assess the extent to which acquired religious representations may be automated.

<sup>16</sup> Rose, Buckwalter & Turri (2014) claim that their data (concerning the ascription of beliefs to people who profess delusional beliefs but fail to act consistently with these beliefs) indicates that delusions are “stereotypical beliefs”, because frequent assertion is a very much more powerful cue to belief ascription than behavior. If they are right, then – given that they satisfy the frequent assertion condition – prompts are stereotypical beliefs too. Citing comparative reaction time data, Rose et al. also argue that delusions are not atypical beliefs, or beliefs that people ascribe only reluctantly. It is important to recognize, however, that delusions might count as genuine, even stereotypical, beliefs without being full-strength beliefs. Beliefs that drive frequent assertion but not other behavior may be recognized as genuine and common beliefs and yet the folk might recognize a distinction between them and beliefs that drive behavior more broadly. Rose et al. did not include the crucial condition which would bear on this question: a within-subjects condition whereby participants were asked to compare assertion without other behavior to assertion coupled with other behavior.

## References

- Almaric, M. & Dehaene, S. 2016. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Science* 113: 4909–4917.
- Astuti, R. & Harris, P.L. 2008. Understanding morality and the life of the ancestors in rural Madagascar. *Cognitive Science* 32: 713-740.
- Atran, S. 2002. *In gods we trust: The evolutionary landscape of religion*. New York, NY: Oxford University Press.
- Atran, S. & Henrich, J. 2010. The Evolution of Religion: How Cognitive By-Products, Adaptive Learning Heuristics, Ritual Displays, and Group Competition Generate Deep Commitments to Prosocial Religions. *Biological Theory* 5: 18-30.
- Barkow, J.H., Cosmides, L., & Tooby, J. (eds.). 1992. *The adapted mind: Evolutionary psychology and the generation of culture*. New York, NY: Oxford University Press.
- Barlev, M., Mermelstein, S. & German, T. 2017. Core intuitions about persons coexist and interfere with acquired Christian beliefs about God. *Cognitive Science* 41: 425-454.
- Barlev, M., Mermelstein, S. & German, T. In press. Representational coexistence in the God concept: Core knowledge intuitions of God as a person are not revised by Christian theology despite lifelong experience. *Psychonomic Bulletin & Review*. DOI: <https://doi.org/10.3758/s13423-017-1421-6>
- Barrett, J. L., & F. Keil. 1996. Conceptualizing a Non-Natural Entity: Anthropomorphism in God Concepts. *Cognitive Psychology* 31: 219–47.

Barrett, J.L. 1999. Theological Correctness: Cognitive Constraint and the Study of Religion. *Method & Theory in the Study of Religion* 11: 325–339.

Barrett, J.L. 2004. *Why would anyone believe in God?* New York, NY: AltaMira Press.

Barrett, J.L. & Trigg, R. 2016. Cognitive and evolutionary studies of religion. In Roger Trigg & Justin L. Barrett (eds.) *The Roots of Religion: Exploring the Cognitive Science of Religion* (1-15). London: Routledge.

Bayne, T. 2010. Delusions as Doxastic States: Contexts, Compartments, and Commitments. *Philosophy, Psychiatry, and Psychology* 17: 329–336.

Bayne, T. & Pacherie, E. 2005: In Defence of the Doxastic Conception of Delusion. *Mind and Language* 20: 163-188.

Bering, J. M. 2002. Intuitive conceptions of dead agents' minds: The natural foundations of afterlife beliefs as phenomenological boundary. *Journal of Cognition and Culture* 2: 263-308.

Bloom, P. 2004. *Descartes' baby: How child development explains what makes us human*. London: Arrow Books.

Boyer, P. 2000. Functional origins of religious concepts: Ontological and strategic selection in evolved minds. *Journal of the Royal Anthropological Institute* 6: 195–214.

Boyer, P. 2008. Religion: Bound to believe? *Nature* 455: 1038–1039.

Boyer, P., and C. Ramble. 2001. Cognitive templates for religious concepts: Cross-cultural evidence for recall of counter-intuitive representations. *Cognitive Science* 25: 535–564.

Chudek, M., McNamara, R.A., Birch, S., Bloom, P. & Henrich, J. In press. Do minds switch bodies? Dualist interpretations across ages and societies. *Religion, Brain & Behavior*, forthcoming.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. 1999. Sources of mathematical thinking: Behavior and brain-imaging evidence. *Science* 284: 970–974.

Dennett, D.C. 1969. *Content and consciousness*. London: Routledge.

Dennett, D. C. 1991. *Consciousness Explained*. New York: Little, Brown.

Deutsch, R., Gawronski, B., & Strack, F. 2006. At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology* 91: 385–405.

Duhaime, E. 2015. Is the call to prayer a call to cooperate? A field experiment on the impact of religious salience on prosocial behavior. *Judgment and Decision Making* 10: 593-596.

Edelman, B. 2009. Markets: Red Light States: Who Buys Online Adult Entertainment? *Journal of Economic Perspectives* 23: 209-220.

Feigenson, L., Dehaene, S. & Spelke, E. 2004. Core Systems of Number. *Trends in Cognitive Sciences* 8: 307-314.

Funkhouser, E. 2017. Beliefs as signals: A new function for belief. *Philosophical Psychology* 30: 809-831.

Gervais, W. M., Willard, A. K., Norenzayan, A., & Henrich, J. 2011. The cultural transmission of faith: Why innate intuitions are necessary, but insufficient, to explain religious belief. *Religion* 41: 389-410.

Greenwald, A.G. & Krieger, L.H. 2006. Implicit Bias: Scientific Foundations. *California Law Review* 94: 945-967

Harris, P., & Giménez, M. 2005. Children's acceptance of conflicting testimony: The case of death. *Journal of Cognition and Culture* 5: 143-164.

Hasson, U., & S. Glucksberg. 2006. Does negation entail affirmation? The case of negated metaphors. *Journal of Pragmatics* 38: 1015-32.

Heider, F. & Simmel, M. 1944. An experimental study of apparent behavior. *The American Journal of Psychology* 57: 243-259.

Hurley, S L. 1998. Vehicles, contents, conceptual structure, and externalism. *Analysis* 58: 1-6.

Jenson, J.C. 2016. The Belief Illusion. *British Journal for the Philosophy of Science* 67: 965-995.

Levy, N. 2015. Neither fish nor fowl: Implicit attitudes as patchy endorsements. *Noûs* 49: 800-823.

Levy, N. 2017a. Religious beliefs are factual beliefs: Content does not correlate with context sensitivity. *Cognition* 161: 109-116.

Levy, N. 2017b. Implicit Bias and Moral Responsibility: Probing the Data. *Philosophy and Phenomenological Research* 94: 3-26.

Levy, N. In press(a). You Meta Believe It. *European Journal of Philosophy*.

Levy, N. In press(b). In Praise of Outsourcing. *Contemporary Pragmatism*.

Malhotra, D. K. 2008. (When) are Religious People Nicer? Religious Salience and the 'Sunday Effect' on Pro-Social Behavior. Harvard Business School NOM Working Paper No. 09-066.

< <http://dx.doi.org/10.2139/ssrn.1297275> >

- Mandelbaum, E. 2016. Attitude, inference, association: On the propositional structure of implicit bias. *Noûs* 50, 3: 629–658.
- McCauley, R. N. 2011. *Why Religion Is Natural and Science Is Not*. New York: Oxford University Press.
- McNamara, T.P. 2005. *Semantic Priming: Perspectives from Memory and Word Recognition*. Hove: Psychology Press.
- Norenzayan, A., S. Atran, J. Faulkner, and M. Schaller. 2006. Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science* 30: 531–553.
- Pina V., Castillo A., Cohen Kadosh R., Fuentes LJ. 2015. Intentional and automatic numerical processing as predictors of mathematical abilities in primary school children. *Frontiers in Psychology* 6: 375.
- Powell, R., & Clarke, S. 2012. Religion as an evolutionary byproduct: A critique of the standard model. *The British Journal for the Philosophy of Science* 63: 457–486.
- Purzycki, B.G. & Willard, A.K. 2016. MCI Theory: A Critical Discussion. *Religion Brain and Behaviour* 6: 207–248.
- Rose, D., Buckwalter, W. & Turri, J. 2014. When Words Speak Louder Than Actions: Delusion, Belief, and the Power of Assertion. *Australasian Journal of Philosophy* 92: 683–700.
- Russell, Y.I. & Gobet, F. 2013. What is Counterintuitive? Religious Cognition and Natural Expectation. *Review of Philosophy and Psychology* 4: 715–749.
- Schwitzgebel, E. 2002. A phenomenal, dispositional account of belief. *Noûs* 36: 249–275.
- Shtulman, A. & Harrington, K. 2016. Tensions between science and intuition across the lifespan. *Topics in Cognitive Science* 8: 118–137.
- Sperber, D. 1996. *Explaining Culture*. Oxford: Blackwell Publishers.
- Sterelny, K. 2003. *Thought in a hostile world*. New York: Blackwell.
- Sterelny, K. 2015. Content, Control and Display: The Natural Origins of Content. *Philosophia* 43: 549–564.
- Sterelny, K. In press. Religion Re-Explained. *Religion, Brain & Behavior*. DOI: <https://doi.org/10.1080/2153599X.2017.1323779>
- Stich, S. 1978. Beliefs and subdoxastic states. *Philosophy of Science* 45: 499–518.
- Van Leeuwen, N. 2014. Religious Credence is not Factual Belief. *Cognition* 133: 698–715.



Xygalatas, D. 2013. Effects of religious setting on cooperative behaviour. A case study from Mauritius. *Religion, Brain and Behavior* 3: 91-102.

Zimmerman, A.Z. 2007. The Nature of Belief. *Journal of Consciousness Studies* 14: 61-82.

Zimmerman, A.Z. 2018. *Belief: A Pragmatic Picture*. Oxford: Oxford University Press.