

In Praise of Outsourcing.

Neil Levy

Department of philosophy, Macquarie University, Sydney 2019, Australia, and Uehro Centre for Practical Ethics, University of Oxford, OX1 1PT, United Kingdom

neil.levy@philosophy.ox.ac.uk

Abstract

What explains the context sensitivity of some (apparent) beliefs? Why, for example, do religious beliefs appear to control behaviour in some contexts but not others? Cases like this are heterogeneous, and we may require a matching heterogeneity of explanations, ranging over their contents, the attitudes of agents and features of the environment. In this paper, I put forward a hypothesis of the last kind. I argue that some beliefs (religious and non-religious) are coupled to cues, which either trigger an internal representation or even partially constitute the beliefs. I show that such coupling will give rise to the context-sensitivity, without entailing that religious believers take a different attitude to belief content.

Keywords

religious belief – cognitive science of religion – mismatch – delusion

1 Introduction

Are religious beliefs the same *kind* of state as beliefs about more mundane matters? For a variety of reasons, some people are tempted to deny that they are. In this paper, I will focus on one motivation for denying that religious beliefs are the same kind of state as more mundane beliefs: their apparent context sensitivity. Whereas ordinary beliefs are apt to guide behavior in any context to which their content is relevant, religious beliefs guide behavior only when they are cued. This difference in functional role (together with other functional differences) suggests that they are a different kind of state to other beliefs.

My primary aim in this paper is to explain the context sensitivity seen with religious beliefs. I will argue that the context sensitivity arises from the manner in which they are *coupled* to external states of affairs. This coupling – and therefore the kind of context-sensitivity it is apt to give rise to – is not unique to religious beliefs: it may be seen in some of our most familiar and most mundane beliefs too. Whether this coupling is a reason to hold that coupled beliefs are a different kind of state to uncoupled is a difficult question. Coupled beliefs exhibit different properties from uncoupled (though there is also heterogeneity among such beliefs) and we may regard these differences as significant enough to justify a distinction in kind. But the differences are fully explained without supposing that agents take a different kind of attitude to the content of their coupled representations.

2 The Context Dependence of Religious Beliefs

There is a range of evidence suggesting that the degree to which people behave consistently with their religious beliefs is context-dependent. Consider, for example, what Malhotra has dubbed the *Sunday Effect*: Christians behave in ways that accord better with the norms they (apparently) accept more on Sundays than on other days of the week. Malhotra found that religious Christians were significantly more likely to respond to an appeal to bid for a charity auction than were nonreligious controls, but only on Sundays. The effect was limited to bidding with a charitable motivation.¹ Edelman found a similar effect of Sundays on internet pornography consumption rates in highly religious US states versus non-religious states: while there was no difference in porn consumption rates overall, rates were significantly lower in religious states than non-religious states on Sundays alone.² Plausibly, the fact that religion is especially salient to Christians on Sundays explains why their behavior accords better with Christian norms on that day alone.

The effect of reminders of religion on behavior is not limited to Christians. Shopkeepers in Morocco gave significantly more to charity while the call to prayer was audible than when it could not be heard.³ Xygalatas found a similar effect of religious versus secular setting on behavior in an economic game among Hindus in Mauritius.⁴ Nor is the effect limited to behavior that accords with moralized norms: making religion salient seems to trigger a different conception of death in both Spanish children⁵ and among the Vezo people of Madagascar.⁶ In both cases, respondents ascribed significantly more of a continuing mental life to a recently deceased person following a religious prime than following a secular prime.

Cases in which belief/behavior match seems to be sensitive to context have inspired a small but growing literature in philosophy and cognitive science. It is widely held these mismatches indicate that the representational state is not a genuine belief, or a belief of the same kind as a factual belief. Thus, for instance, some people have suggested that the belief/behavior mismatch apparently seen in monothematic delusions suggests that

¹ Deepak K. Malhotra, "(When) are Religious People Nicer? Religious Salience and the 'Sunday Effect' on Pro-Social Behavior," Harvard Business School NOM Working Paper No. 09-066, 2008.

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

² Benjamin Edelman, "Red Light States: Who Buys Online Adult Entertainment?" *Journal of Economic Perspectives* 23 (2009): 209-220.

³ Erik P. Duhaime, "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 (2015): 593-596.

⁴ Dimitris Xygalatas, "Effects of religious setting on cooperative behaviour. A case study from Mauritius," *Religion, Brain and Behavior* 3 (2013): 91-102.

⁵ Paul Harris and Marta Giménez, "Children's acceptance of conflicting testimony: The case of death," *Journal of Cognition and Culture* 5 (2005): 143-164.

⁶ Rita Astuti and Paul L. Harris, "Understanding morality and the life of the ancestors in rural Madagascar," *Cognitive Science* 32 (2008): 713-740.

delusions are not beliefs,⁷ while Eric Schwitzgebel has suggested that such cases require us to postulate “in-between beliefs.”⁸ Neil Van Leeuwen has also attempted to draw conclusions about the nature of the attitude from mismatch cases, arguing that religious beliefs have two properties more mundane beliefs lack.⁹ They are “practical setting dependent” – that is, they guide behavior only in some of the contexts to which their contents are relevant – and even when they guide behavior, they do so incompletely. Religious believers operate with a “two-map cognitive structure”,¹⁰ keeping track of both the ordinary facts and of conflicting religious beliefs. This two-map structure enables them to ensure that they do not act in ways that would generate factual evidence that would count against their religious beliefs that as the stakes rise, guidance is by the light of mundane beliefs. They are far more likely to pray for the occurrence of naturalistically explicable events than for violations of the laws of nature,¹¹ for instance. Van Leeuwen cites these (alleged) functional differences between religious and mundane beliefs as evidence that they are different kinds of attitudes.

I will argue that both practical setting dependence and the limitations in the extent to which religious beliefs govern behavior when they are salient are the result of how (some) religious beliefs are coupled to cues and artifacts in the world. In the absence of relevant cues, the representation will not govern behavior. Even in their presence, it may not govern behavior in the full range of ways characteristic of uncoupled beliefs, because some of the representational work is offloaded onto the cue, and the syntactical properties of cues are typically less flexible than the syntactical properties of more purely internal representations.

Before turning to a discussion of the coupling of representations, let me note an important motivation for looking to an alternative explanation for the phenomena Van Leeuwen observes. Explaining these phenomena by reference to the kind of attitudes religious beliefs allegedly are leaves us unable to account for the fact that often enough behavior is guided by religious beliefs in just the same kind of way in which behavior is guided by mundane beliefs. For instance, often believers guide their behavior by their religious beliefs in circumstances in which their factual falsity would predict disaster. This fact provides some motivation for holding that agents take the factual attitude to their contents – the same attitude they take toward mundane beliefs.

⁷ Greg Currie, “Imagination, hallucination and delusion,” *Mind and Language*, 15 (2000): 168-183; Greg Currie and Ian Ravenscroft, *Recreative Minds: Imagination in Philosophy and Psychology* (Oxford: Oxford University Press, 2002).

⁸ Eric Schwitzgebel, “In-between believing,” *Philosophical Quarterly* 51 (2001): 76-82; “Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief,” *Pacific Philosophical Quarterly* 91 (2010): 531-553.

⁹ Neil Van Leeuwen, “Religious Belief is not Factual Belief,” *Cognition* 133 (2014): 698-715.

¹⁰ Neil Van Leeuwen, “Two Paradigms for Religious Representation: The Physicist and the Playground,” *Cognition* 164 (2017): 206-211.

¹¹ Justin L. Barrett, “How Ordinary Cognition Informs Petitionary Prayer,” *Journal of Cognition and Culture* 1 (2001): 259-269.

Consider, for instance, medieval trials by ordeal. When the facts of a case were difficult to ascertain, defendants were sometimes given the choice of performing some task that could be expected to cause them great harm, or confess their guilt. For instance, they might be asked to undergo trial by fire; this might involve them walking a certain distance holding irons that were red-hot. This procedure was held to be a reliable means of ascertaining guilt or innocence, because God would intervene to protect the innocent from harm. To contemporary eyes, trial by ordeal seems not only barbaric, but also useless as a forensic tool. But the (admittedly sparse) surviving records tell a different story.

A surprisingly large proportion of those who underwent trial by ordeal were acquitted; that, in turn, indicates that they were not badly burned. Leeson¹² suggests that priestly manipulation, rather than divine intervention, explains this fact (he points out that the rules governing the trials allowed plenty of opportunity for such manipulation). Why would priests manipulate the heat of the iron or the water? Leeson suggests that the behavior of the accused carried a great deal of information about their guilt. If they were reluctant, that was evidence that they greatly feared the ordeal, and – given that they believed that the innocent would be spared – that was in turn evidence of their guilt (Leeson suggests, the guilty would usually avoid the trial, especially since pleading guilty often carried a lesser penalty than being found guilty by trial; many people did indeed settle before the trial). But if they were not reluctant, this was evidence that they were innocent, and the priest might therefore intervene to ensure that they did not suffer much.

The important point for our purposes is that this mechanism works only if the accused really – factually – believe that their innocence is protective. It is only if this is the case that their behavior carries information about their guilt or innocence. In these cases, it seems that religious beliefs governed behavior when they should not have, if these beliefs guide behavior in a manner that is constrained by factual beliefs with conflicting contents. Ordinary believers seem typically to have really believed that their innocence was protective. But what about the priests? Doesn't their manipulation require that they do *not* believe that God will protect the innocent? It is at least possible that priests would generally have a more sophisticated faith than their parishioners, since they would have more opportunity to observe God's strange inactivity. That doesn't mean that they therefore would lose faith in the ability of God to protect the innocent. They might instead see themselves as carrying out God's work when they manipulated trials, Leeson suggests.¹³

¹² Peter T. Leeson, "Ordeals," *Journal of Law and Economics* 55 (2012): 691–714.

¹³ Mightn't we explain the behavior of ordinary believers in a similar way, by attributing to them a sophisticated faith, according to which God works through priestly manipulation? The problem with this hypothesis is that it doesn't explain those cases in which people settled rather than face the ordeal: knowing that one only has unflinchingly to go through with the trial to be acquitted, because willingness to undergo it would lead to priestly manipulation, should ensure that almost everyone was acquitted. Note, too, that we cannot explain the pattern of results by suggesting that the priests spared those they knew to be innocent: trial by ordeal was reserved for cases in which there was a paucity of evidence.

Trial by ordeal is not evidence against the practical setting dependence of religious beliefs. The trials were conducted by priests and were preceded by multiple rituals. The belief *that God will protect me from the fire* might have guided behavior only because religious primes were highly salient to the believer. But it is evidence that *when* they guide behavior, sometimes religious beliefs do so in precisely the same way as beliefs with more mundane contents. On Van Leeuwen's account, the person who undergoes trial by ordeal keeps track of mundane facts as well as sometimes conflicting religious beliefs. Since they factually believe *that the fire will burn me*, even if they also religiously believe *that God will protect me*, we ought to see evidence of great fear from them. But, apparently, their religious belief was sufficiently strong to ensure that the exhibited much less fear than those who were guilty. It is only if religious beliefs are factual beliefs, or close kin to such beliefs, that this mechanism works.

Admittedly, some of the cases in which agents are (apparently) caused by their religious beliefs to engage in behaviors that they might have avoided if they were instead guided by conflicting factual beliefs can be explained consistently with the two map cognitive structure framework. Behavior consistent with the religious belief may be explained as costly signalling of group allegiance,¹⁴ or credibility enhancing displays.¹⁵ When the costs are especially high, the signal is especially clear.¹⁶ This hypothesis suggests an alternative explanation of the data Leeson cites. On this explanation, the ordeal would not be a means of detecting guilt; rather it would be a means of distinguishing the more from the less committed.

However, while this kind of account can explain some cases in which religious beliefs are not constrained by factual belief, there are others which it leaves mysterious. There is wide agreement that whatever the original function (if any) of religious beliefs, they have come to play a role in stabilizing groups which are too large for individuals to be able to track reputations. If religious beliefs are to play that role, however, they must guide behavior broadly. More pointedly, guidance must extend not only to contexts in which there is a conflict between the behavior that accords with the religious belief and the behavior that better accords with mundane beliefs, but also to contexts in which there is no opportunity to signal to others, because behavior goes unobserved. Religious beliefs promote prosocial behavior toward in-group members, and thereby enable large scale societies, when they postulate the existence of omniscient gods¹⁷ who sanction

¹⁴ Richard Sosis, "Religious behaviors, badges, and bans: Signaling theory and the evolution of religion," In *Where God and Science Meet: How Brain and Evolutionary Studies Alter Our Understanding of Religion, Volume 1: Evolution, Genes, and the Religious Brain*, ed. Patrick McNamara (Westport, CT: Praeger Publishers, 2006), pp. 61–86.

¹⁵ Joseph Henrich, "The evolution of costly displays, cooperation and religion," *Evolution and Human Behavior* 30 (2009): 244-260.

¹⁶ I am grateful for Neil Van Leeuwen for discussion of this point.

¹⁷ Or, if not omniscient, at least gods that are "full access social strategic agents," who are omniscient with regard to moral behaviors. See Pascal Boyer, *Religion Explained: The Evolutionary Origins of Religious Thought* (New York: Basic Books, 2001), pp. 156-7.

immoral behavior.¹⁸ High Gods enable cooperation in anonymous societies by taking on the role of observing behavior: *they* observe when *we* cannot. Now, for High Gods to enable cooperation, they must cause agents regularly to act in ways that conflict with how they would act were they not under the gaze of the supernatural watcher. High Gods must cause agents to resist temptation to defect on cooperative arrangements when defection is in their interests. And that entails that High Gods lead agents to act in ways that accord with their religious beliefs when doing so incurs a significant cost, compared to acting in ways that are guided by their (supposedly persisting) factual beliefs.¹⁹

Given that religious beliefs often appear to guide behavior in the same sort of way as beliefs with more mundane contents, we should be wary of suggestions that they are attitudes of a distinct kind. An account which explains the practical-setting dependence (at least) sometimes observed while respecting the ways in which they sometimes occupy the full set of functional roles of more mundane beliefs, as well as the ways in which they sometimes occupy a narrower set of functional roles, should be preferred to one that postulates novel attitudes.

I will suggest that the context-sensitivity exhibited by religious beliefs arises because they belong to the class of *coupled representations*. Coupled representations are coupled to cues that may serve as triggers for them, and which sometimes even serve as partial constituents of the belief. Agents token coupled representations when the cues are available to play a triggering or constitutive role; in their absence, they do not token the representation and their behavior is therefore not guided by it.

3 Coupling Representations

The functional role of representations is behavior guidance. Such guidance requires a map, and the representation provides such a map. We need to represent spatial relations, causal relations and temporal relations. Very often, however, representations do not need to be available except in the presence of cues that indicate that they are needed. For most purposes, a map of downtown that I can consult only when there are indications I am headed there is pretty much as good as one that I can consult all the time. Such a map could be coupled to these indications.

¹⁸ Ara Norenzayan, *Big Gods: How Religion Transformed Cooperation and Conflict* (Princeton: Princeton University Press, 2013).

¹⁹ The High Gods hypothesis is controversial. It is worth noting, however, that the argument of this paper does not depend on its central claim (that the emergence of High Gods is required for the development of large-scale societies) being true. It is enough for my purposes if High Gods religions can promote prosociality, and that the causal mechanism implicates surveillance. Even some of the harshest critics of the high gods hypothesis as an account of the origin of large-scale societies concede that they can promote cooperation (see Russell D. Gray and Joseph Watts, "Cultural macroevolution matters," *Proceedings of the National Academy of Sciences* 114 (2017): 7846-7852). I thank Thomas Coleman for prompting me to think about this point more deeply.

These considerations suggest that coupling representations to indicators may not have significant costs. In fact, there is reason to think that such coupling will often be adaptive. Evolution is sensitive, in the long run, to very small differences in costs. Time and energy are costs, so it is often adaptive to trade off accuracy for speed or frugality: hence the pervasiveness of quick and dirty heuristics when slower and more demanding cognitive strategies might track truths more accurately. One way to save on processing resources and time is to cut down on the number and the richness of internal representations when there are adequate alternatives available. There is extensive evidence that we actually do this.²⁰ Our internal models of the world are less detailed than we tend to think. Change blindness experiments provide evidence for this claim: We may fail to notice large changes in the world when we expect it to be stable,²¹ because we do not have a detailed internal model of the world to compare it to. As a result of our lack of a detailed internal representation, we may even fail to notice that the person we have been talking to has been substituted by another who shares only the grossest similarities to the original,²² and fail to notice the substitution of every word of a text presented electronically with gibberish if the substitution is timed so that we are always reading the only genuine words on the screen.²³

The visual saccades in which participants in many experiments engage also provides evidence that they are retrieving visual information from the environment as and when it is needed, and that in turn indicates that their representation of the visual scene is partially constituted by the visual scene itself.²⁴ This offloading of representational duties onto the very world represented is highly adaptive. Whereas trading off speed for accuracy is all things considered adaptive because the benefits outweigh the costs, when we offload representations onto the world there is no loss of accuracy. On the contrary, the world is always a more accurate model of itself than any internal representation could be: it updates in real time and captures every detail of itself. Possible costs might arise if perceptual access to the world is too unreliable, or in the form of increased access time, if we can process an internal model more quickly than an external one. But these costs, should they arise, may be outweighed by the gains in accuracy and the savings that come from obviating the need to devote resources to the maintenance of an internal representation.²⁵

²⁰ Andy Clark, *Being There: Putting Brain, Body, and World Together Again* (Cambridge, Mass.: The MIT Press, 1997).

²¹ Daniel J. Simons and Daniel T. Levin, "Change Blindness," *Trends in Cognitive Sciences* 1 (1997): 261-267.

²² Daniel J. Simons and Daniel T. Levin, "Failure to Detect Changes to People During a Real-World Interaction," *Psychonomic Bulletin & Review* 5 (1998): 644-649.

²³ Keith Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin* 124 (1998): 372-422.

²⁴ Andy Clark, *Supersizing the Mind* (Oxford: Oxford University Press, 2008).

²⁵ Some organisms – like us – want their maps to be available for offline manipulation and planning. This kind of broader availability might be achieved by uncoupling representations from indicators that they are required. But it can also be achieved by expanding the range of indicators that cue map availability. If there are regularly recurring circumstances in which we need our maps to be available, we can cue

Given that the world is its own best model, and utilizing it as such is efficient, the coupling of representations to the world might be expected to evolve under favorable conditions. Centrally, the organism's perceptual link to the world must be sufficiently reliable, such that the external information is available to shape cognition and action when it is needed. If the organism's perceptual link to these properties is highly reliable, then why go to the extra trouble of constructing an internal representation when the world is right there to play that role?

Representations may be coupled to what I called above *indicators*, which are cues that they are needed. In cases in which the representation is coupled to the world as its own model, however, the cue may not serve merely as an indicator. Its role may extend beyond merely triggering the representation, to helping to *constitute* it. In cases like this, my representation of a visual scene is itself partially external to me. Call a cue which partially constitutes a representation a *prop*. Cases in which a cue serves as a trigger for a purely internal representation might be compared to the way in which smartphones can key reminders to locations using GPS or cell tower data. The cue serves as a trigger, but once the reminder is triggered the representational load is carried entirely internally (to the phone or the mind). Outsourcing with props might be better compared to the manner in which we use maps on our smartphones, relying on the map to represent not only the immediate area, but also (via the blue dot for example) our current location and direction of travel. Perhaps we do better to see the location-based reminder model and the blue dot model as points on a continuum rather than as wholly distinct. Representational load is a continuous quantum (cases in which there are *no* internal representations at all are probably non-existent) and triggers for relevance may be internal or external and perhaps partially one and partially the other.

Given that coupling of representations to indicators and props is likely to be adaptive, we should expect to observe coupling. There is indeed evidence that some representations are coupled. Importantly, the coupling of representations to cues may give rise to context effects. This kind of context-sensitivity may arise in at least two ways for coupled representations. First, behavior may fail to be guided by a coupled representation that is relevant to the context when an indicator for relevance is absent (or in some way masked). It may be obvious to an observer that a professed belief of the agent is relevant in the context, but because the indicator is absent, the representation is not triggered and it therefore fails to play its characteristic role. Second, when behavior is driven by such a representation we may see it track changes in the world without giving rise to a distinctive phenomenology or to awareness that behavior is tracking such changes: because the agent lacks a representation of the property to which behavior is cued – because, that is, she relies on the world to represent the property for her – she may not be conscious of the change in the property, but her behavior may track it nevertheless. In that case, too, we may see context effects which may not track agents' professed beliefs. In fact, there is empirical evidence for just such context dependence: in experiments that require motor responses, agents are capable of correcting for surprisingly large changes in the placement of a target without

availability to indicators of those circumstances. Creatures like us, who engage in domain-general cognition, may even be able to be self-cueing.

conscious awareness of such changes.²⁶ Because the cue to which the representation is coupled alters, actions guided by the representation alters too, but because the person has only coarse-grained internal representations they fail to notice either the change or the motor adjustment.

We should expect to see this kind of lack of awareness seen in the motor adjustments in the experiments just mentioned in cases in which representations are coupled to props. In these cases, our internal representations may not contain representations of the prop itself – because it serves as its own representation – and we may therefore fail to notice changes in the prop. When cues serve as indicators and not props, they may themselves be internally represented. Even in those cases, we may not notice any alterations in them: the very stability that makes them apt for coupling entails that they are unlikely to be salient to us. Attention is itself a limited resource and will be deployed selectively to properties it is adaptive to track. In cases in which we couple representations to save on representational costs, the property that takes the load is expected to be highly stable and therefore may fade into the taken-for-granted background. Because these properties are unattended, changes in them typically escape detection, and we may fail to notice the ways in which our own behavior alters in response. It is worth noting here that when religious beliefs fail to guide behavior in ways we might expect, believers seem unaware of the apparent discrepancy.²⁷

It is time to put some flesh on the bones of the hypothesis. In the next section, I will begin to sketch more concretely how coupled representations actually function and to provide some examples. Before I provide some specifically religious examples, I will discuss a non-religious case that may involve coupled representations: delusional belief.

4 Coupled Representations in Action.

Sufferers from monothematic delusions profess to believe some bizarre proposition (*that my wife has been replaced by an impostor; that I am dead; that the image I see when I stand in front of a mirror is a stranger who is living in my home*). Notoriously, the behavior of sufferers may fail to be consistent with those claims. For example, sufferers from Capgras delusion (the delusion that someone close to the patient has been replaced by a replica) often fail to be troubled by the substitution. They rarely report the impostor to police or express concern for their real wife or husband. Like those who fail to act consistently with the religious beliefs, moreover, sufferers typically fail to notice the apparent inconsistency between what they profess to believe and how

²⁶ Bruce Bridgeman, Marc Kirch and Alan Sperling, “Segregation of cognitive and motor aspects of visual function using induced motion,” *Perception and Psychophysics* 29 (1981): 336–42; Melvyn Goodale, Denis Péliesson and Claude Prablanc, “Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement,” *Nature* 320 (1986): 748–50.

²⁷ Pascal Boyer, “Why “belief” is hard work: Implications of Tanya Luhrmann’s *When God talks back*,” *Hau: Journal of Ethnographic Theory* 3 (2013): 349-57; Nicolas Baumard and Pascal Boyer, “Religious beliefs as reflective elaborations on intuitions: A modified dual-process model,” *Current Directions in Psychological Science* 22 (2013): 295-300.

they often act. It is almost as if they operate a two-map cognitive structure, with their delusional beliefs constrained by competing factual beliefs.

Delusional patients may even acquire and lose the delusion from context to context. Consider FE, who suffers from mirrored-self misidentification: the delusion that the image he sees when he stands in front of a mirror is another person. Breen and colleagues report the following exchange between one of the authors and FE, which occurred while he was positioned so that he could see both the examiner's reflection and his own in a mirror:

Examiner: Who does it look like? Have you seen this person in here before? (*pointing to the reflection of the examiner*).

FE: That's you.

Examiner: That's me?

FE: Yes.

Examiner: Me, here? (*pointing to herself*) What's my name?

FE: I don't know, oh yes, it's Nora.

Examiner: Nora, that's right. So that's me in the mirror?

FE: Yes.

Examiner: That's my reflection?

FE: Yes.

Examiner: And who is that? (*pointing to FE's reflection*).

FE: I don't know what you would call him. It makes me a bit sick because he moves about freely with us. I don't be too friendly because I don't see it does him any good.²⁸

FE understands what a mirror is and what a reflection is. But the knowledge controls his behavior only when the reflection is not his own. When he sees his own reflection, this knowledge seems to become inactive. He doesn't appear to experience himself as changing his mind, though; he fails to notice the discrepancy between his fluent deployment of the mirror as a mirror and his claim that the person in the mirror is a stranger.

Other cases are if anything even more dramatic. DS experienced Capgras delusion for his father, but the delusion remitted when they spoke on the telephone: only the visual percept provoked the delusion.²⁹ Anosognosia for hemiplegia – denial of paralysis on one side of the body – sometimes remits temporarily in response to a very low-tech treatment: vestibular stimulation, which is simply the introduction of cold water into the ear canal.³⁰ After stimulation, patients sometimes frankly acknowledge the

²⁸ Nora Breen, Diane Caine, Max Coltheart, Julie Hendy and Corrine Roberts, "Towards an Understanding of Delusions of Misidentification: Four Case Studies," *Mind and Language* 15 (2000): 74-110; at 84-5.

²⁹ William Hirstein and Vilayanur Ramachandran, "Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons," *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264 (1997): 437-444.

³⁰ Stefano Cappa, Roberto Sterzi, Giuseppe Vallar and Edoardo Bisiach, "Remission Of Hemineglect And Anosognosia During Vestibular Stimulation," *Neuropsychologia* 25 (1987): 775-82.

paralysis, but a few hours later revert to denial. There is no apparent phenomenology, or even acknowledgement, of a change of mind.

Delusions may usefully be understood, I suggest, as arising from the way in which relevant beliefs arise from coupling. Consider, for illustration, the Capgras delusion. While the details of just how it is involved in the generation of a representation like *my wife has been replaced* are controversial, it is widely agreed that a perceptual abnormality is central to the syndrome. The Capgras delusion arises when one of the perceptual pathways involved in person recognition is lesioned. The pathway underlying conscious visual perception is intact (hence the patient has no trouble in recognizing that the familiar person *appears to be* someone they know), but the visual pathway underlying a feeling of familiarity is dysfunctional. The full-blown representation of the person, however, is coupled to the person herself: perception of the person highly reliably generates a feeling of familiarity that is a constituent of the representation of the person. When the feeling is lacking, only a partial representation is triggered. They see person as qualitatively (but not numerically) identical to their family member.

On this hypothesis, there is very little in the way of inference required for the patient to conclude that the person they see is an impostor (contra standard two factor theories of delusions). Rather than deducing that the person is an impostor on the basis of the lack of feeling of familiarity, the patient simply fails to token their normal representation of the person. The system is set up such that perception of the person is supposed to trigger the relevant representation via the feeling of familiarity; its breakdown here leads to the representation being partially tokened. The patient may best be described as *seeing that* the person is not who they seem to be. There may be further (subpersonal) inference involved, to the conclusion *that's an impostor* (it is noteworthy that the precise content of the delusion is a function of cultural narratives: the impostor might be an alien or a robot today). But there is no inference drawn to the conclusion (say) *that's not my wife*.

For this hypothesis to be viable, the feeling of familiarity must play some role in cognition beyond helping to trigger beliefs like this one. Coupled representations are coupled to cues that are reliably available. In general, the cue is available at all for some independent reason. The feeling of familiarity presumably evolved to enable us to discriminate between cases, and playing this discriminatory role requires that in its absence our classifications would fall below some sufficiently high threshold of reliability. Plainly, the feeling of familiarity did not evolve in order to play a role in discriminating between intimates and impostors. Our face perception is reliable enough without supplementation by this feeling (in fact, were we not highly reliable at identifying intimates on the basis of visual cues *other than* the feeling of familiarity, the Capgras delusion would not arise at all: patients would at least sometimes be unsure of the identity of their intimate, rather than confident that they are an impostor). The feeling of familiarity must therefore have as its proper function discriminating between other kinds of cases. Whatever the story that should be told here, its highly reliable availability in cases like the ones in which we perceive an intimate allows it to play a role in helping to trigger the representation.

Delusions arise, I suggest, when the cues to which certain representations are coupled (because these properties are extremely stable features of the contexts in which the representation has the function of guiding behavior) fail to be instantiated. In these cases, the person may fail to token the full representation that would be rational in the context. But because neither coupling itself nor the cues to which representations are coupled are salient, they typically fail to notice that anything is amiss. The person fails to notice that there is a discrepancy between the behavior they exhibit when in the grip of the delusion and the behavior that they exhibit at other times.

If this explanation is roughly on the right track, then delusions arise from a deficit in coupling: we token the full representation of a familiar person only when she elicits the feeling of familiarity in us. This would be a limiting case of coupling, inasmuch as here the cue for eliciting the representation is the person herself. Because the cue is the very person who is represented, it will take gross pathology to disrupt this kind of coupling. If, as is more standard, the cue is something that is reliably, but not invariably associated with the representation, we will see failures of coupling without pathology: when the cue is not instantiated when it should be.

The kinds of context-dependence we see in religious belief cases may arise as a result of a similar sort of outsourcing to cues. Religious beliefs may be easily disrupted when the cue is not instantiated. In many cases, at least, the coupling of religious representations is a recent exaptation, by cultural evolution, of the machinery of coupling, and the representations are coupled to cultural cues and artifacts which are far less stable than the feeling of familiarity that accompanies normal person perception.

It is controversial whether religion is an adaption or a byproduct of adaptations. It is, however, widely held that religious dispositions build on our mentalizing capacities and perhaps on our hair trigger disposition to detect agency in the world.³¹ These dispositions are adaptations for navigating the social world and for responding to threats and opportunities, not for religion, but they provide a fertile ground for the development of religious thoughts and behaviors, which may or may not themselves have come to be a target for subsequent natural selection.

These religious dispositions in turn provide fertile ground for the development of High Gods religions. These newer religions are cultural adaptations which build on the foundations provided by an earlier-emerging religious cognition, in a manner similar to how that earlier-emerging cognition built on a suite of cognitive dispositions that evolved for other functions. High Gods religions also build on dispositions with independent evolutionary histories; in particular, our disposition to cooperate when under the watchful gaze of other agents. This disposition to cooperate when we are observed is obviously adaptive: not only do we risk retaliation if we are seen to defect, but maintaining a good reputation for cooperative behavior is extremely important in the small bands in which the first Homo Sapiens lived. As a result, “watched people are

³¹ Justin L. Barrett, “Exploring the natural foundations of religion,” *Trends in Cognitive Sciences* 4 (2000): 29-34; Justin L. Barrett, *Why Would Anyone Believe in God?* (Lanham, MD: AltaMira Press, 2004).

nice people”.³² The degree to which we are prosocial is in fact exquisitely sensitive to cues of gazes. We orient automatically to faces,³³ and even schematic representations of eyes increase our disposition to cooperate.³⁴ The disposition automatically to detect eyes (and to modulate behavior accordingly) is, like the disposition to detect agency, hyperactive; because false negatives are on average more costly than false positives, the mechanisms are easily triggered by stimuli which bear even a passing resemblance to their targets. As a result of the facts that the disposition to detect eyes is so sensitive and that detecting eyes has effects on behavior, cultural evolution has coupled a disposition to prosocial behavior and representations associated with such behavior to images of eyes and other cues suggestive of vigilant, powerful, agents.

Images and other reminders of eyes are almost ubiquitous in cultures with High Gods.³⁵ Even the theological understanding of High Gods as located above us (as the name suggests), and the metaphors used to capture their relationship to believers, reinforce the thought that they are watching their flock (as the shepherd does his). Because the disposition to prosocial behavior (more precisely, the disposition to prosocial behavior in excess of the behavior to which we are disposed in virtue of the building blocks of morality, like reciprocal altruism and kin selection, which enable the cooperative behavior seen in human cultures prior to the emergence of large scale societies) is driven by coupled representations, however, and the representation is coupled to cues that are not omnipresent, the disposition itself may be lost and gained in response to cues suggesting the presence or the gaze of the gods.³⁶

³² Norenzayan, p. 19.

³³ Jon Driver, Greg Davis, Paola Ricciardelli, Polly Kidd, Emma Maxwell and Simon Baron-Cohen, “Gaze Perception Triggers Reflexive Visuospatial Orienting,” *Visual Cognition* 6 (1999): 509-540.

³⁴ Mary Rigdon, Keiko Ishii, Motoki Watabe and Shinobu Kitayama, “Minimal social cues in the dictator game,” *Journal of Economic Psychology* 30 (2009): 358-367; Norenzayan, pp. 19-23 discusses multiple further examples.

³⁵ Norenzayan, pp. 23-29.

³⁶ While a number of studies have appeared to indicate that images of watching eyes increased prosocial behavior, other have failed to replicate it (for instance, Adam M. Sparks, and Pat Barclay, “No effect on condemnation of short or long exposure to eye images,” *Letters on Evolutionary Behavioral Science* 6 (2015): 13-16). Two recent meta-analyses have shown that the effect size is not significantly different from zero (Stefanie B. Northover, William C. Pedersen, Adam B. Cohen and Paul W. Andrews, “Artificial surveillance cues do not increase generosity: Two meta-analyses,” *Evolution and Human Behavior* 38 (2017): 144-153). However, the hypothesis presented here does not depend on the claim that such cues increase prosocial behavior. Rather, it depends on the claim that specifically religious primes increase prosociality (at least with regard to fellow group members) and that claim is strongly supported (Luke Galen, “Does religious belief promote prosociality? A critical examination,” *Psychological Bulletin* 138 (2012): 876-906; Azim F. Shariff, Aiyana K. Willard, Teresa Andersen and Ara Norenzayan, “Religious Priming: A Meta-analysis with a Focus on Prosociality,” *Personality and Social Psychology Review* 20 (2016): 27-48). Some authors have suggested that this effect is best explained by a belief that behavior is observed by a divine being (Roy F. Baumeister, Isabelle M. Bauer and Stuart A. Lloyd, “Choice, free will, and religion,” *Psychology of Religion and Spirituality* 2 (2010): 67-

Unsurprisingly, then, many of the cases in which we see shifts in the behavior of religious believers driven by shifts in context concern prosocial behavior (recall the Sunday effect, for instance). There are, however, cases which do not concern this kind of behavior. Van Leeuwen for instance, is impressed by the evidence that believers seem to switch between a religious conception of death, according to which some mental processes continue afterwards, and a more purely biological conception, depending on context. Nevertheless, this shift, too, may arise from coupled representations, or (better) from a kind of shuffling between coupled and uncoupled representations. There is data to suggest that the differing conceptions of death between which the respondents in these experiments oscillate have different developmental trajectories, with the biological conception emerging first and subsequently co-existing with the later developing religious conception.³⁷ Thus young children are unlikely to attribute mental states to the dead at all, but as they grow older these attributions become more frequent. These attributions occur, however, only given appropriate – external – cues. That suggests that the biological representation of death is an uncoupled representation (or a representation coupled to the mere invocation of death, which serves as trigger), to which we tend to revert in the absence of cues for the religious representation. The competing religious representation is coupled to a narrower range of indicators, as a result of the agent's learning history.³⁸

Religious representations may be coupled not just to indicators but also to props, and such coupling might explain the ways in which they sometimes seem to depart from the full functional profile of beliefs when they are triggered. In many cultures, people sacrifice foodstuffs to the gods, but eat the food themselves.³⁹ This seems an instance in which behavior is governed by a state that falls short of a genuine belief: if they believe that the food is for the gods, they should not consume it, and if they consume it, they should come to be aware that there is no genuine sense in which it is sacrificed. We may explain the apparent incongruity in their behavior by recognizing that coupled representations may be patchy representations: patchy in virtue of the fact that their internal vehicles are not structured in the rich kinds of ways seen with representations (coupled to indicators or uncoupled) that are fully internally realized. The props take on

82). It seems likely, however, that the modality need not be visual or supernatural. Indeed, secular primes like *court* have the same effect on behavior. This fact suggests that the coupling of representations to cues may be as effective in non-religious societies, in which civil institutions take over the role of promoting cooperation.

³⁷ Paul Harris, *Trusting What You're Told: How Children Learn from Others* (Cambridge, Mass.: Harvard University Press, 2012).

³⁸ Note that this hypothesis is consistent with the claim that dualism is innate (e.g. Paul Bloom, *Descartes' Baby* (New York: Basic Books, 2004): on at least some ways of cashing out innateness, it is compatible with requiring external input. For instance, a trait might reliably develop across a very broad range of environments, and it might be species typical, and yet require specific external input for its development. So long as that external input is itself reliably found in the organism's environment, the trait may be buffered against perturbation. On this view, what is purely internal is a *disposition* to develop certain concepts, given the right inputs. These inputs may then come to serve as indicators for tokening the representations

³⁹ Boyer 2001.

some of the representational roles played by more purely internal state. Because the props, which themselves have few or no syntactic properties, carry some of the load, the representation itself may not be rich enough to entail the full range of inferences that an internal representation with the same content would have had. Little wonder agents fail to detect incongruity between their actions and their beliefs: there may in fact *be* no such incongruity. Incongruity would require a far richer internal representation. There are cases and cases here, depending on the richness of the internal representations and the entailments that such a belief would have were it fully internally realized.

Plausibly, the religious domain is not a natural kind.⁴⁰ That being the case, religious beliefs may have relatively little in common with each other, beyond the facts about content or causes that leads us to categorize them as religious. Some religious beliefs may be uncoupled representations, while some are coupled to indicators and some to props. It may even be the case that religious beliefs with similar contents are coupled to props in one context and indicators in another. It is instructive here to contrast the Sunday effect with the effects of the call to prayer. The Sunday effect appears to persist across much of the day. It continues to drive behavior long after Church services have ended. In contrast, the call to prayer appears to have very short lived effects on behavior: though 100% of respondents chose the most charitable option when the call to prayer was audible, less than 50% of those who responded in the 20 minutes following the call chose this option.⁴¹ While the numbers of participants who responded in this time interval was too low for us to be very confident that the effect dissipated as rapidly as that suggests, the contrast with Sunday effect is striking. This contrast suggests that the call to prayer served as a prop of the religious representations of Duhaime's participants, whereas American Christians have their similar representations coupled to indicators.

Speculatively, the relative frequency of reminders of gods might predict whether relevant beliefs are coupled to props or to indicators. For American Christians living in a largely secular society, reminders of religion are less frequent than in some other contexts. That is, props that could sustain religious beliefs are sparse in the environment. That being the case, religious representations may persist only to the extent to which they are more fully internalized once triggered. In contrast, the call to prayer sounds 5 times a day, everyday. It is therefore available to take up the representational slack to a much greater extent. Its availability enables Muslims who live and work in places in which it is audible to offload representational duties. Note that this need not mark a distinction between Islam and Christianity. In medieval Europe, for instance, the Church bells and the visible steeple might allow for even greater offloading of representational duties than we see in contemporary Marrakesh. In general, given that we can expect offloading of representational duties to props whenever the world is available to serve as its own best model, we should expect that to the extent to which religious cues are reliably available in agents' environment, they will have sparser internal representations. As a consequence, such representations will have a sparser set of functional roles once the prop is removed. We may also see much more rapid decay rates for propped representations after removal, on the plausible assumption that richer representations are likely to persist for longer durations.

⁴⁰ Ibid.

⁴¹ Duhaime 2015.

This speculation gives rise to a testable prediction: we ought to see differences in the duration of guidance by religious cues across otherwise similar populations as a function of internalisation of their religious representations, where representations will tend to be internalised to a greater extent in contexts in which religion is less pervasive in public life. The United States, in which the presence of religion varies from state to state (and especially from north to south) provides an ideal context for testing this prediction. Holding self-reported religious belief and denomination fixed, we ought to see faster decaying religious representations in Southern Christians than Northern Christians.⁴²

5 Conclusion

Religious beliefs seem to have some puzzling features. They sometimes exhibit a surprising sensitivity to context, failing to govern behavior in many situations to which their content seems relevant. Even when they do play a role in behavior, they sometimes fail to play the full set of functional roles we expect from beliefs. As a result of these two facts, religious believers may behave in ways that appear to conflict with their sincere professions of faith. Yet at other times, their behavior may suggest an unflinching and wholehearted belief. I have suggested that we can explain all these features of religious beliefs by supposing that they are coupled representations: coupled to worldly cues that trigger or even sometimes partially constitute them.

When religious representations are coupled to indicators, they will not govern behavior in contexts to which their content is relevant if those contexts fail to feature their indicators. In such cases, agents' behavior will be governed by competing representations. When they are coupled to props, they may lack the full functional roles we associate with beliefs in the absence of the prop, or because subpersonal expectation of the prop leads to a relatively sparse representation. Sometimes religious beliefs are full-strength beliefs, whether because they have all the internal structure of such beliefs or because they are coupled to props that compensate for their sparseness. Sometimes they lack sufficient internal structure to occupy the full range of functional roles we expect of beliefs.

Whether we should reserve the label 'belief' to states that have the full range of functional roles as those represented internally or not is a difficult question. It is not a question I will try to settle here. I have tried to show that we can explain the ways in which functional roles may be attenuated by reference to how representations may be coupled. It is worth remarking, in conclusion, that because coupling (to indicators or props) is not something seen only with religious beliefs (and perhaps not always with religious beliefs), whatever taxonomy of beliefs we end up adopting will not have religious beliefs on one side and mundane beliefs on the other.

Acknowledgements

I am grateful to Thomas J. Coleman III and a reviewer for this journal (who identified himself as Konrad Talmont-Kaminski) for very helpful comments on this paper.

⁴² I owe this suggestion to Thomas Coleman.

References

- Astuti, Rita and Paul L. Harris, P.L. 2008. Understanding morality and the life of the ancestors in rural Madagascar. *Cognitive Science* 32: 713-740.
- Barrett, Justin L. 2000. Exploring the natural foundations of religion. *Trends in Cognitive Sciences* 4: 29-34.
- Barrett, Justin L. 2001. How Ordinary Cognition Informs Petitionary Prayer. *Journal of Cognition and Culture* 1: 259–269.
- Barrett, Justin L. 2004. *Why Would Anyone Believe in God?* Lanham, MD: AltaMira Press.
- Baumard, Nicolas and Pascal Boyer. 2013. Religious beliefs as reflective elaborations on intuitions: A modified dual-process model. *Current Directions in Psychological Science* 22: 295-300.
- Baumeister, Roy F., Isabelle M. Bauer and Stuart A. Lloyd. 2010. Choice, free will, and religion. *Psychology of Religion and Spirituality* 2: 67–82
- Bloom, Paul. 2004 *Descartes' Baby*. New York: Basic Books.
- Boyer, Pascal. 2001. *Religion Explained: The Evolutionary Origins of Religious Thought*. New York: Basic Books.
- Boyer, Pascal. 2013. Why “belief” is hard work: Implications of Tanya Luhmann’s *When God talks back*. *Hau: Journal of Ethnographic Theory* 3: 349-57.
- Breen, Nora, Diane Caine, Max Coltheart, Julie Hendy and Corrine Roberts. 2000. Towards an Understanding of Delusions of Misidentification: Four Case Studies. *Mind and Language* 15: 74-110.
- Bridgeman, Bruce, Marc Kirch, and Alan Sperling, A. 1981. Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics* 29: 336–42.
- Cappa Stefano, Roberto Sterzi, Giuseppe Vallar and Edoardo Bisiach. 1987. Remission Of Hemineglect And Anosognosia During Vestibular Stimulation. *Neuropsychologia* 25: 775–82.
- Clark, Andy. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, Mass.: The MIT Press.
- Clark, Andy. 2008. *Supersizing the Mind*. Oxford: Oxford University Press.
- Currie, Greg. 2000. Imagination, hallucination and delusion. *Mind and Language* 15: 168-183.

- Currie, Greg and Ian Ravenscroft, I. 2002. *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford: Oxford University Press.
- Driver Jon., Greg Davis, Paola Ricciardelli, Polly Kidd, Emma Maxwell and Simon Baron-Cohen. 1999. Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition* 6: 509-540.
- Duhaime, Erik P. 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, Benjamin. 2009. Red Light States: Who Buys Online Adult Entertainment?" *Journal of Economic Perspectives* 23: 209-220.
- Galen, Luke, W. 2012. Does religious belief promote prosociality? A critical examination. *Psychological Bulletin* 138: 876 -906.
- Goodale, Melvyn, Denis Pélisson and Claude Prablanc. 1986. Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* 320: 748-50.
- Gray, Russell. D. and Joseph Watts. 2017. Cultural macroevolution matters. *Proceedings of the National Academy of Sciences* 114: 7846-7852.
- Harris, Paul and Marta Giménez. 2005. Children's acceptance of conflicting testimony: The case of death. *Journal of Cognition and Culture* 5: 143-164.
- Harris, Paul. 2012. *Trusting What You're Told: How Children Learn from Others*. Cambridge, Mass.: Harvard University Press.
- Henrich, Joseph. 2009. The evolution of costly displays, cooperation and religion. *Evolution and Human Behavior* 30: 244-260.
- Hirstein, William and Vilayanur Ramachandran. 1997. Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 264: 437-444.
- Leeson, Peter T. 2012. Ordeals. *Journal of Law and Economics* 55: 691-714.
- Malhotra, Deepak 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> >
- Norenzayan, Ara. 2013. *Big Gods: How Religion Transformed Cooperation and Conflict*. Princeton: Princeton University Press.

- Northover, Stefanie B., William C. Pedersen, Adam B. Cohen and Paul W. Andrews. 2017. Artificial surveillance cues do not increase generosity: Two meta-analyses. *Evolution and Human Behavior* 38: 144-153.
- Rayner, Keith. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124: 372-422.
- Rigdon, Mary, Keiko Ishii, Motoki Watabe and Shinobu Kitayama 2009. Minimal social cues in the dictator game. *Journal of Economic Psychology* 30: 358–367.
- Schwitzgebel, Eric. 2001. In-between believing. *Philosophical Quarterly*, 51(202), 76-82.
- Schwitzgebel, Eric. 2010. Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, 91(4), 531-553.
- Shariff, Azim F., Aiyana K. Willard, Teresa Andersen and Ara Norenzayan, 2016. Religious Priming: A Meta-analysis with a Focus on Prosociality *Personality and Social Psychology Review* 20: 27-48.
- Simons, Daniel J. and Daniel T. Levin. 1997. Change Blindness. *Trends in Cognitive Sciences* 1: 261-267.
- Simons, Daniel J. and Daniel T. Levin. 1998. Failure to Detect Changes to People During a Real-World Interaction. *Psychonomic Bulletin & Review* 5: 644-649.
- Sosis, Richard. 2006. Religious behaviors, badges, and bans: Signaling theory and the evolution of religion. In *Where God and Science Meet: How Brain and Evolutionary Studies Alter Our Understanding of Religion, Volume 1: Evolution, Genes, and the Religious Brain*, ed. P. McNamara (Westport, CT: Praeger Publishers), pp. 61–86.
- Sparks, Adam M. and Pat Barclay. 2015. No effect on condemnation of short or long exposure to eye images. *Letters on Evolutionary Behavioral Science* 6: 13–16.
- Van Leeuwen, Neil. 2014. Religious Belief is not Factual Belief. *Cognition* 133: 698-715.
- Van Leeuwen, Neil. 2017. Two Paradigms for Religious Representation: The Physicist and the Playground. *Cognition* 164: 206-211.
- Xygalatas, Dimitris. 2013. Effects of religious setting on cooperative behaviour. A case study from Mauritius. *Religion, Brain and Behavior* 3: 91-102.