

What is Counterintuitive? Religious Cognition and Natural Expectation

Yvan I. Russell · Fernand Gobet

© Springer Science+Business Media Dordrecht 2013

Abstract What is ‘counterintuitive’? There is general agreement that it refers to a violation of previously held knowledge, but the precise definition seems to vary with every author and study. The aim of this paper is to deconstruct the notion of ‘counterintuitive’ and provide a more philosophically rigorous definition congruent with the history of psychology, recent experimental work in ‘minimally counterintuitive’ concepts, the science vs. religion debate, and the developmental and evolutionary background of human beings. We conclude that previous definitions of counterintuitiveness have been flawed and did not resolve the conflict between a believer’s conception of the supernatural entity (an atypical “real kind”) and the non-believer’s conception (empty name/fictional). Furthermore, too much emphasis has been placed on the universality and (presumed) innateness of intuitive concepts (and hence the criteria for what is counterintuitive)—and far too little attention paid to learning and expertise. We argue that many putatively universal concepts are not innate, but mostly learned and defeasible—part of a religious believer’s repertoire of expert knowledge. Nonetheless, the results from empirical studies about the memorability of counterintuitive concepts have been convincing and it is difficult to improve on existing designs and

Yvan Russell was partly funded by the German Initiative of Excellence of the German Science Foundation (DFG) and by the EC-FP6 EXREL project grant 43225.

Yvan Russell wrote this manuscript as part of the Explaining Religion (EXREL) Project and the British Academy Centenary Research Project.

Y. I. Russell
Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford, UK

Y. I. Russell (✉)
Department of Psychology, Middlesex University, London, UK
e-mail: yvanrussell@gmail.com

Y. I. Russell
Junior Research Group “Evolution of Cooperation and Prosocial Behavior”, Courant Research Centre “Evolution of Social Behaviour”, University of Göttingen, Göttingen, Germany

Y. I. Russell · F. Gobet
Department of Psychology, Brunel University, Uxbridge, UK

F. Gobet
Department of Psychological Sciences, University of Liverpool, Liverpool, UK

methodologies. However, future studies in counterintuitive concepts need to embed their work in research about context effects, typicality, the psychology of learning and expertise (for example, the formation of expert templates and range defaults), with more attention to the sources of knowledge (direct and indirect knowledge) and a better idea of what ‘default’ knowledge really is.

1 Introduction

What does it mean to say that something is counterintuitive? Recent papers have transformed the word ‘counterintuitive’ into a technical definition about the nature of religious concepts (e.g., Barrett 2008). The key issue is about breaking the rules of a core concept, but keeping that core concept in place. A “dog who talks in the English language” is counterintuitive because the ability to talk is a violation of our knowledge about dogs. If we accept that such a dog really exists without having seen it in person, then we are forcing ourselves to mentally hold together two incompatible ideas (think of holding two magnets together of the same charge, where the magnets slip from side to side and resist being held together). Being forced to accept a counterintuitive concept feels this way. People generally seek to preserve their pre-existing beliefs (Taber and Lodge 2006; Upal 2011a); and because the category of DOG is firmly established in the mind as a non-talking animal, we might—upon first hearing about the talking dog—*search for alternative explanations* of the statement. Here are some options:

- (1) The core concept has been mislabelled (it is not a dog).
- (2) The core concept is correct, but the behaviour is not really attached to it (the talking came from elsewhere).
- (3) The event is being reported as a counterfactual (a ‘what-if’ scenario, such as in a fairy tale).
- (4) Someone is trying to deceive you (by convincing you that such a thing exists).
- (5) The core concept needs to be updated (dogs can actually speak English, although you have not witnessed it personally).

There is a sixth option. You can retain the mystery by not explaining the conflict at all, but making an attempt to accommodate or justify the new information in the context of what you already know (Pyysiäinen 2003; Upal 2011a). Consider the following example from Nigeria about witches:

“According to local beliefs, witches have four eyes which allow them to see in the dark and to see the spirits that walk the earth. It is also believed that witches can split their body and soul in two, allowing them to commit nefarious deeds while the physical body remains asleep” (Lainé 2007, p. 15).

The witch is a real-life person who wears outlandish clothes and facial paint (see photo in Lainé 2007, p. 15), but the supernatural dimension is mere hearsay. Yet, in every world religion, there are typical sets of *counterintuitive* claims, consisting of religious concepts that “typically comprise claims or statements that violate people’s ideas of what commonly takes place in their environment” (Boyer 1994, p. 35; also see Atran and Norenzayan 2004; Boyer 1993, 2001; Boyer and Ramble 2001; Pyysiäinen 2002, 2003; Steadman et al. 2009; Tweney 2012; Upal 2010, 2011a).

The aim of this paper is to deconstruct the notion of ‘counterintuitive’, the meaning of which remains unclear (Guthrie 2007; Näreaho 2008) despite some attempts to define it

precisely (e.g. Barrett 2008; and papers reviewed below). Researchers in this area fall into two main camps (Upal 2010): the *concept* view (putting emphasis on the concepts themselves as the source of counterintuitiveness) and the *context* view (putting emphasis on the information that preceded counterintuitiveness). We suspect that writers in the concept school have tended to overstate the universality of their postulated concepts. Below, we investigate the underlying theoretical issues behind counterintuitiveness, how it relates to our evolved human cognition, our cultural backgrounds, and to the structural properties of the concepts themselves. We provide examples of concepts which are counterintuitive (and which are not), examine how we gather information about concepts in the first place (asking questions about naturalness), examine the role of context, and critique the recent line of research about “minimally counterintuitive” (MCI) concepts. Throughout this article, we put special emphasis on *religious* cognition—because much recent research in counterintuitiveness has been conducted with the specific aim of explaining religious concepts (e.g. Barrett 2008).

2 Mind, Concept, and Culture

We start by looking at religious beliefs as examples of counterintuitive cognition. Why do people believe in supernatural concepts? If we study the formation, maintenance, and perpetuation of counterintuitive beliefs, we can ask three inter-related questions about how these beliefs arose:

- (1) Do our minds actively shape the religious and cultural information into *typical forms* that fit our evolved cognitive biases, with respect to information processing and memory storage?
- (2) Are certain packets of information—*due to their internal structure*—more contagious and transmissible, compared to other packets of information?
- (3) To what extent does surrounding cultural knowledge determine the structure of a person’s religious beliefs (perhaps overriding natural biases)?

Looking at the first question, we ask about the constraints of the human mind. “What we believe about gods, spirits, and ancestors is firmly constrained by what we can encode, process, and recall” (Whitehouse 2001, p. 169). This viewpoint comes from the “cognitive science of religion” (CSR), a cognitive approach to studying religion (for background, see Atran 2006; Atran and Norenzayan 2004; Barrett 1999, 2007a, b; Boyer 1994, 2000, 2001; Cohen et al. 2008; Guthrie 1980, 1993, 2007; Laidlaw 2007; Legare et al. 2009; McCauley 2000; Norenzayan et al. 2006; Pyysiäinen 2003, 2004, 2013; Slone 2005; Tweney 2012; Whitehouse 1992, 1995, 2001, 2004a, b, 2008; for critiques of CSR, see Bloch 2008; Cho 2013; Day 2007; Guthrie 2007; Laidlaw 2007; Näreaho 2008; Nynäs 2008; Salazar 2010, and Saler 2010). In contrast to the “psychology of religion”—which has traditionally studied the effect of religion *on* people (Gorsuch 1988; Strunk 1971)—CSR studies the effect of people *on* religion. For example, Guthrie (1980, 1993), amongst others, has argued that anthropomorphism (a ubiquitous human bias, see Barrett 2011; Waytz et al. 2010) is a major cognitive predisposition that is conducive to religious belief. Subsequent empirical studies have supported this idea in adults and children (Atran and Norenzayan 2004; Barrett 2007b; Barrett and Keil 1996; Barrett and Richert 2003). The relevant message of CSR is that certain religious notions are widespread *because* they comprise ideas that dovetail particularly

well with evolved human cognition (see Atran and Norenzayan 2004; Norenzayan et al. 2006; Whitehouse 2001, 2004a).

For the second question, we exit the brain and consider only the internal semantic structure of the concepts that our society has given us. As Hampton (2010) says, “concepts are not just personal mental entities, but have a life of their own in a social community” (p. 306). Religion is a distributed phenomenon (Whitehouse 2001, 2004a), consisting of a human population who share a set of mental representations which are relatively homogeneous (cf. Hampton 2010; Ross and Tidwell 2010; Sperber 1985; Whitehouse 2004a, and Upal 2011a). The cultural epidemiology approach (Sperber 1985) is analogous to how medical epidemiologists study the spread and distribution of diseases (for critiques of this approach, see Guthrie 2007; Laidlaw 2007; Slone 2005; Upal 2010; Whitehouse 2001; cf. Hampton 2010), but instead of tracking diseases, it is tracking mental representations (cf. Cullen 2000). Unlike diseases, however, mental representations do not pass directly from one body to another. In order for a mental representation to be passed from one person to another, the hearer needs to actively reconstruct the informative intention of another person (Whitehouse 2001; cf. Sperber and Wilson 1996). When an idea is reconstructed in another hearer’s head, the hearer needs to understand the *structural property* of that concept. Consider the aforementioned Nigerian witch. As shown in Fig. 1 (this diagram is inspired by the view of Boyer 2000), the core of the WITCH concept is WOMAN, with all standard assumptions about women attached. This core concept is what anchors your comprehension. Then, you find two deviations from the standard assumptions (i.e. having four eyes, and departing her body during sleep). According to recent research (reviewed below), the structural property of a concept is a predictor of how successfully the concept will spread in a population (e.g. see Boyer 2000, 2001; Boyer and Ramble 2001; Norenzayan et al. 2006). Certain concepts appear to be “cognitively optimum”—easy to remember and pass on—and consequently they become widespread in a population (Barrett 2007b; Boyer 1994, 2001; Whitehouse 2004a).

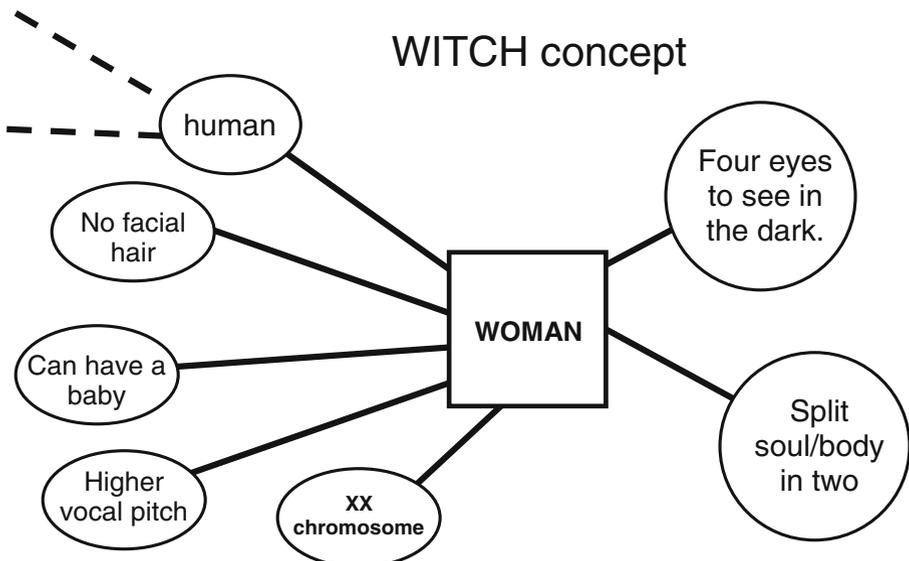


Fig. 1 Diagram of WITCH concept

The third question is about cultural background (Upal 2010). People learn religious and supernatural ideas through their social and cultural context (Boyer 1994, 2001; Pyysiäinen 2003, 2004; Steadman et al. 2009; Upal 2010, 2011a; Whitehouse 1992). Culture is an information source, where the teachings, priorities, and attitudes of others provide a *template* for how to organize your thoughts (Colby and Cole 1973; Ross and Tidwell 2010). People develop “habits of the mind” (Ross and Tidwell 2010) that structure their understandings of the world. Religions are conceptual frameworks for interpreting the world (Atran and Norenzayan 2004; Boyer and Ramble 2001; Guthrie 1980; Hick 1983; Koltko-Rivera 2004; McCauley 2000; Pyysiäinen 2003; Steadman et al. 2009; Upal 2010, 2011a; Vidal and Kleeberg 2007; Whitehouse 1992; cf. Ross and Tidwell 2010). Different religious traditions vary immensely in the content of their teachings (see entries in Hinnells 1984), but all religions feature four main types of repertoire that reference the supernatural (Boyer 1994): (1) existence of non-observable entities, (2) supernatural status of human individuals, (3) causal links between the supernatural and observed events, and (4) knowledge of supernatural episodes (cf. Atran and Norenzayan 2004). As Hampton (2010) says, “concepts are part of a broader causal-explanatory scheme with which we understand the world” (p. 305). In a religious tradition, people develop a world view that serves as an interpretive lens on the world (Atran and Norenzayan 2004; Koltko-Rivera 2004; but see Whitehouse 2004a, p. 50–51). Believers may interpret the clearly visible world as the handiwork of supernatural agents (Grim 2006; Upal 2010). They may interpret the “fleeting shadow... seen in a clearing” (Boyer 1994, p. 92) as an encounter with a ghost (see Boyer 2001; Guthrie 1980, 1993). Most importantly, they listen to testimony about things unseen (Audi 2002; Goldberg 2010; Whitehouse 2004a; cf. Ganea et al. 2007, and Hampton 2010). People routinely augment personal observations with testimony from others (see Goldberg 2010; Hampton 2010; Sommerfeld et al. 2007).

Why does a person believe in counterintuitive witches? As outlined above, this question may have a tripartite answer: (1) evolved cognitive biases (our minds shape the way that concepts are formulated), (2) inherent structural properties of concepts (some structures are catchier than others), and (3) cultural context (we are persuaded by surrounding cultural opinion). The three explanations are not incompatible with each other, and all three should be considered when we study counter-intuition (i.e. as three types of independent variable). Before going further, we need to clarify how “counterintuitive” is defined.

3 What are Counterintuitive Concepts?

New information is always put into context. Sometimes, it is integrated quite easily into what a person already knows. For example, when a chess master encounters a novel configuration of chess pieces, the pattern is immediately recognized as being similar to basic structures (chunks and templates) of previously learned patterns (Gobet and Simon 1996, 1998; Gobet 1998; we return to chess examples later in this article). Other newly encountered information is “troublesome” (a word used by educational psychologists, e.g. Meyer and Land 2005): difficult to learn because it is too alien (cf. Hampton 2010). In the colloquial sense, “counterintuitive” simply refers to new information that clashes with prior knowledge (Pyysiäinen 2004). The eminent psychologist Leon Festinger (1957) explored such clashes in his “theory of cognitive dissonance”. When a new piece of information appears on your mental horizon, it is assessed in one of three ways:

dissonant, consonant, or irrelevant to what you already know. As Festinger defines dissonance: “x and y are dissonant if not-x follows from y” (1957, p. 13).

What is the “-intuitive” being countered when you say “counterintuitive”? It seems a misnomer at first, because the standard psychology definition of “intuitive” refers to a *process*, not a belief. For example, Betsch (2008) defines intuition as a “process of thinking” (p. 3) that occurs without conscious awareness, whereby knowledge in long term memory feeds into “a *feeling* that can serve as a basis for judgments and decisions” (p. 3, italics added; also see Gobet and Chassy 2009; Chassy and Gobet 2011). In other words, Betsch (2008) identified intuition as the *process* leading to a “hunch”. “Hunches” are often characterized as the quick-but-fallible initial conclusion that a person makes, in contrast to slower and more analytic forms of decision making (e.g. see Berger 2013, about judicial hunches). Those who study CSR use the word “intuitive” differently. They refer to *intuitive ontology* (Atran and Norenzayan 2004; Barrett 2008; Boyer 1994, 2000, 2001): beliefs about characteristics of entities based on everyday “folk” experience (i.e. collective representations by a group; cf. Hampton 2010). Beliefs can be described as *reflective* (consciously known) or *non-reflective* (believed implicitly) (Barrett 2007b; cf. Berger 2013). A person could not possibly cogitate on every possible unseen combination of features in the environment (e.g., dogs cannot fly, pigs cannot fly, cats cannot fly, etc.) because there are an infinite number of propositions available. At any given moment, cognition is usually ad hoc, largely shaped by the specific features of the environment or ecological niche (Barrett 2011; Clark 1998). According to the CSR view, many ontological expectations are “intuitive” because they are non-reflective (Barrett 2007b), formed though the process of intuition (*sensu* Betsch 2008), based on the shorthand of category membership, resulting in a belief that is not necessarily brought into explicit attention (Barrett 2007b; Boyer 1994, 2000, 2001)—until it is countered.

To clarify the CSR definition of “counterintuitive”, consider the following statements. Which of the following are counterintuitive?

- (1) A man who has always appeared calm suddenly erupts in anger.
- (2) An elderly woman in a wheelchair jumps out of her wheelchair and then stands on her head.
- (3) A flipped coin comes up heads twelve times in a row.
- (4) One hundred coins were thrown onto the floor and then (without human intervention), the coins spelled out the name “Charles Darwin”.
- (5) A cat has been seen floating on air.
- (6) A cat has been described as dog-shaped.
- (7) A cat has been described as weighing 5,000 kg.
- (8) A witch has four eyes to see the spirit world and can split her soul in two in order to leave her body at night (Lainé 2007).

Many of the above situations are unusual, but only statements (5) and (8) are properly counterintuitive (*sensu* Barrett 2008). Statements (1) to (3) are surprising but *not* counterintuitive because they describe entirely plausible events. The calm man (statement 1) had never been observed to get angry, but you can easily alter your impression of the man’s personality (he was not as calm as he looked; he gets angry in certain situations, etc.). For the elderly woman (statement 2), the most rational explanation is that she was only *pretending* to need a wheelchair. A coin flipping onto heads twelve times in a row (statement 3) has a probability of .00024 but it is still entirely possible. Statement (4)

represents a grey area. Although it does not transgress the laws of physics for the carelessly thrown coins to spell out “Charles Darwin”, the probability is so vanishingly small that philosophers call this kind of event a “quasi-miracle” (Dodd 2011): de facto impossible even if it can theoretically happen (see Dodd 2011, for a summary of this debate).

In CSR, the word “counterintuitive” has a very specific meaning. It does not simply refer to the dissonant, absurd, unlikely, impossible, or unfamiliar. Here, the word “counterintuitive” is used within an interpretive framework for analyzing the transmission of religious ideas across populations (Atran 2006; Atran and Norenzayan 2004; Barrett 2008; Barrett and Nyhof 2001; Boyer 1994, 2001; Boyer and Ramble 2001; Guthrie 2007; Pyysiäinen 2002, 2004; Slone 2005; Tweney 2012; Whitehouse 2004a,b; for critiques of counterintuitiveness research, see Bloch 2005, 2008; Cho 2013; Day 2007; Näreaho 2008; Nynäs 2008, and Salazar 2010). However, despite this focus on religion, there are no compelling reasons cited that differentiate the cognitive mechanisms that occur in *religious* domains and *non-religious* domains (Pyysiäinen 2003, 2013): the content may be different, but the process is the same. When people categorize the supernatural, they are using the same cognitive processes that they use to categorize the mundane (see Atran and Norenzayan 2004) and, indeed, most religious behaviour is just ordinary, everyday, behaviour, but which is overlaid with a supernatural interpretation (Bloch 2008; Steadman et al. 2009). Therefore, in the line of research cited below, it is not uncommon to see *non-religious* counterintuitive concepts (see examples in Table 1) being used to draw inferences about *religious* counterintuitive concepts (e.g. Boyer 2001, pp. 58–104). Religion utilizes two types of contradiction (Boyer and Ramble 2001; Barrett 2008): a “breach” of the normal characteristic (e.g. an invisible horse), and a “transfer” of characteristics from one entity to another (e.g. a “flying horse” has a skill transferred from a bird). In this area of study, Barrett (2008) provided a comprehensive guide, wherein he claimed that “counterintuitive” refers to concepts whose descriptions violate “maturationally natural” expectations (McCauley 2000) about the concept (cf. Barrett et al. 2009; Carey 2011, and Spelke and Kinzler 2007; for a critique, see Upal 2010). These violations (see Barrett 2008) could be *spatial* (space and time), *biological* (contradicting requirements of living things), regarding *animacy* (whether or not self-propelled), *mentality* (whether it is thinking), and *universal* (regarding temporal and continuous existence).

Defined this way, counterintuitive events seem tantamount to “miracles” (see discussion in Pyysiäinen 2004). Hume (1748/1993) defined miracles as violations of the *laws of nature* whose reports contradict the weight of a person’s past experience. However, we need to distinguish between the actual laws of nature and a person’s construal of them (Pyysiäinen 2004). In a scientific framework, if a law of nature is “violated”, then the relevant theory must be substantially updated (or even rejected altogether) in order to assimilate the new information (Dunbar 1995; Popper 1968; Tweney et al. 2006, pp. 494–495; cf. Upal 2010). This is different from supernatural violations (i.e. miracles), which *do not change* the status of the intuitive domains (for background, see Barrett 2008; Pyysiäinen 2004; Upal 2010; and see Boyer 1994, on “credal status”). In the minds of believers throughout history, miracles are “exceptions to the rule” engineered by sentient supernatural beings (Pawlikowski 2007). As Tweney et al. (2006) describe it, “believers [are] asked to accommodate to a fixed, unchanging, even eternal, order” (p. 495; cf. Pyysiäinen 2003). In science, by contrast, “natural laws” are universal statements about observed reality, in the form of “there-is”—and conversely—“there-is-not” statements

Table 1 Selected examples of devised counterintuitive stimuli from empirical studies

Citation	Example(s) of counterintuitive concept	Comment
Barrett and Nyhof (2001)	“a being that can see or hear things no matter where they are. For example, it could make out the letters on a page in a book hundreds of miles away and the line of sight is completely obstructed.” (p. 79).	At least two violations of core concept of human-like being.
Boyer and Ramble (2001)	“There was a person who was at two places at same time. He was at one place and at another place at the same moment”; “There was one person who had no shadow. Even in broad daylight, he did not have a shadow” (p. 561).	One violation to core concept of human being.
Norenzayan et al. (2006)	“Thirsty door”; “Confused table”; “Four-legged student”; “Drying comment” (p. 537)	Examples of minimally counterintuitive
Tweney et al. (2006)	“Giggling Seaweed”, “Melting Lady”, “Running Dying Stone”, “Walking Solidifying Pens” (p. 487).	Two minimally and two maximally counterintuitive
Slone et al. (2007)	“chanting grazing TV”; “teaching lustful macaroni”, “eating religious belief” (p. 361).	Noun as core concept with 1–2 violations
Harmon-Vukić and Slone (2009)	“The pigeon/was in the park on sunny day/ rested on a bench/watch children throw stones on the ground./Smiled at the kids/saw a child complaining/melted in the heat” (p. 62).	Counterintuitiveness occurs due to lack of connection between elements of the story
Gregory and Barrett (2009)	“A fly that is immortal”; “A camel that is invisible”; “A cup that is thinking”; “A potato that is everywhere”; “A cactus that is jumping” (p. 296).	Examples of minimally counterintuitive concepts
Johnson et al. (2010)	“A grass covered mountain that likes to sing to the birds that occupy it when it’s happy and becomes invisible when it’s sad”; “An intelligent Lioness that speaks English, gives birth to human infants, can travel back in time and can transform itself into different animal forms” (p. 115).	Two examples of core concepts (mountain, lioness) with 2 and 4 violations, respectively.
Upal (2011b)	“A villager who was made out of steel met a man from the neighboring village who had wings that he used to fly like a bird.” (p.46); “A villager who could bend spoons with his eyes met a woman from the neighboring village who had ten heads.” (p. 47).	Human concepts with one violation each.
Harmon-Vukić et al. (2012)	“One woman floated in space, could spit fire, and breathe under water.” (p. 139).	Example of maximally counterintuitive concept.

(Popper 1968). Natural laws allow some events to occur, and disallow others. For example, Newton's law of universal gravitation stipulates the conditions whereupon one body in space would fall towards the other (Hawley and Holcomb 2005), *allowing* a smaller body (e.g. a meteor) to fall towards a larger body (e.g. Earth) if the latter is large enough and close enough; and *prohibiting* the smaller object from flying away (unless there is a competing gravitational pull). Regarding the counterintuitive, we can use Barrett's framework to look for "there-is-not" statements (sensu Popper 1968; cf. Atran and Norenzayan 2004, and Boyer 1994, 2001). A human *cannot* be in two places at once (spatial violation). A human *cannot* exist only on Tuesdays (time violation; universal violation). A human *cannot* survive without eating (biological violation). A *statue* of a human being *cannot* move by itself (animacy violation). A statue *cannot* think (mentality violation). Anything that violates "there-is-not" (sensu Popper 1968) in these aforementioned domains is counterintuitive. A crucial point is that *natural laws* refer to conditions that can exist independently of human knowledge (Popper 1968; McCauley 2000), even if we cannot measure or observe them. A person's "feelings of conviction" (Popper 1968, p. 44) might be utterly irrelevant to natural laws (Dunbar 1995; Pyysiäinen 2003, 2004). Therefore, we should conclude that "counterintuitive" refers strictly to the *experiential* inside-the-head world of humans (cf. Barrett 2008; Hampton 2010, and Pyysiäinen 2004): it is an evaluation made by a person, intuitively cross checked against a lifetime of learning (for discussion, see Boyer 1994; also Hume 1748/1993; Festinger 1957; Pyysiäinen 2003; Upal 2010). We defend this position in the remainder of this paper.

Returning to our list, we see that a "cat that floats on air" (statement 5) *would be* counterintuitive by Barrett's (2008) definition (a spatial and biological violation of our knowledge of cats, mammals, and large-bodied animals in general). However, in some settings, even a floating cat can be explained rationally. Floating would not be counterintuitive if the cat lived on the International Space Station. You could explain a floating cat easily because the context (being in outer space) is where mammals can float without trying. The floating cat is counterintuitive *if and only if* the cat is earth-bound (compare to work of Gregory and Barrett 2009). Moving onto statement 6 ("dog-shaped cat"), we find a pure example of dissonance (sensu Festinger 1957) because the "dog shape" is dissonant with "cat shape". However, a "dog-shaped cat" is not an ontological violation and therefore would *not* qualify as counterintuitive (see Barrett 2008). Instead, it is *counterschematic* (Barrett 2008), merely a contradiction in terms (also see Boyer 1994, on "nonschematic"; cf. Akman 2009; Gregory and Barrett 2009, and Johnson et al. 2010). By the same token, the cat that weighs 5,000 kg (statement 7) is bizarre but *not* counterintuitive (Barrett 2008; cf. Barrett and Nyhof 2001). In fact, the 5,000 kg cat might be called counterschematic, because our category of "cat" does not permit even the largest cats to weigh 5,000 kg; yet, it is possible for other mammals (e.g. whales) to weigh that much.

Finally, we come to the Nigerian witch (statement 8). This is a truly counterintuitive concept (sensu Barrett 2008) because the folk concept (woman) is not usually attached to such magical abilities. However, we might ask a difficult question. How does the *actual* comprehension experience differ between statement (8) and statement (1)? The calm man (statement 1) had the capacity to become violently angry, even though that capacity had not been previously witnessed. Similarly, the witch (statement 8) is purported to have abilities that have not actually been witnessed by anybody (unless you believe the rumors). The sudden anger of a calm man (statement 1) may have violated a social impression, but it does *not* violate ontology. Obviously, intuitive ontology references naturalness—and a counterintuitive event represents a violation of that natural state. How do you decide when a concept is 'natural'?

4 Are Some Concepts More ‘Natural’ than Others?

What is a concept? There have been many attempts to answer this question over the years (and here we cannot provide an exhaustive review of concepts that will please everybody), but we can generally identify concepts as being *relational*—either denoting real life phenomenon, or relating to other concepts (for diverse definitions of concepts, see Anderson 1980; Barsalou 2009; Carey 2011; Dennett 1995; Feldman 2003; Hampton 1999, 2010; Keil 1989; Malt et al. 2011; Medin and Rips 2005; Millikan 2009; Murphy 2010; Palmer 1981; Popper 1968). Concepts can occupy different levels of explanations according to the complexity of the referent: at the more abstract level (e.g. GOVERNMENT), a concept might comprise relevant multiple ideas joined together. At a simpler level (e.g. APPLE), they might be considered ‘atomic’ (not reducible into smaller joined ideas) (cf. Anderson 1980; Atran 1998; Feldman 2003). Concepts occur in many forms (Medin and Rips 2005; Murphy 2010): *taxonomic* (e.g. FISH, can swim, has gills, lays eggs, etc.), *thematic* (e.g. DOG, leash, walk, park), *event-based* (e.g. WEDDING, bride, groom, music, etc.), *script-based* (e.g. sequence of events for dining at a restaurant), describing *relations* (e.g. fish LARGER THAN bacteria), or it can be completely ad hoc (e.g. “things to bring on a night fishing trip”, Medin and Rips 2005; cf. Ross and Tidwell 2010, on cultural necessity; cf. Hampton 2010). Often, concepts are difficult to describe in a single word and need modifiers to be added to a core word (Malt et al. 2011).

Concepts have at least two main functions (see Medin and Rips 2005; Murphy 2010): to communicate ideas and as a vehicle for *categories* of referents (which can include everything from concrete objects like TABLE or abstract concepts like JUSTICE). The latter function—categorization—is of particular importance to the study of counterintuitive concepts (see Boyer 1994, 2001). There have been a diversity of schools of thought on what constitutes a “category” (Boyer 1994; Hampton 2010; Keil 1989; Medin and Rips 2005; Millikan 2009; Murphy 2010) but there is general agreement that a category consists of a core concept connected to a list of attributes (as in our FISH category above; cf. Boyer 2000, and Fig. 1). Categorical schemes tend to be hierarchically structured—e.g. dachshund is a kind of dog, dog is a kind of mammal, etc. (Palmer 1981; cf. Atran 1998)—but categories are of limited usefulness if they are either too *broad* (e.g. “all objects”) or too *narrow* (i.e. where the criterion exhausts the category) (Murphy 2010). Ultimately, categories are all imperfect because of their fuzzy boundaries, their inconsistent prototypes, and their opaque definitions (Hampton 2010; cf. Keil 1989, and Millikan 2009). “Categories are useful”, says Murphy (2010), “to the degree that they tell us about their members” (p. 15).

Human children learn to represent the world in at least five domains (Spelke and Kinzler 2007): *object representation* (spatiotemporal principles of cohesion, continuity, and contact), *agents and their actions* (attention to cues of agency, causation, and goals), *number* (differentiations of amounts), *space* (the geometry between objects in the environment), and *social representation* (differentiating and representation people and categories of people). Each domain is learned as a set of regularities, upon which strict constraints are imposed (cf. Nynäs 2008). Can the constraints ever be broken? It depends on what constraint you mean. Akman (2009), a semanticist, identified three types of constraint: (1) “necessary constraint” (necessary to maintain a classification system; their violation is equivalent to counterschematic), (2) “nomic constraint” (perceived laws of nature; their violation is equivalent to counterintuitive), and (3) “conventional constraint” (cultural norms; their violations are “rule breaking”). Two out of

three of these constraints are *relative* (not universal). For example, people will differ in what they call “necessary constraints” (because different people are exposed to different classification schemes) and “conventional constraints” (because cultural conventions are diverse). However—when facing “nomic constraints”—all people should be the same if the laws of physics are the same. A judgment about “counterintuitive” pertains only to “nomic” constraints and when they have been broken (compare to Barrett 2008).

Are some categories universal across all cultures? This is a longstanding question among psychologists and philosophers (see Keil 1989). People seem to agree that the world consists of *natural kinds* (occurring without human agency), *nominal kinds* (artificially defined categories), and *artifacts* (objects), but there is much debate on how to define and differentiate these “kind” categories (Keil 1989; cf. Carey 2011 and Millikan 2009). Onto what “kinds” can we apply counterintuitive violations? Barrett’s (2008) framework clearly stipulates that counterintuitiveness is the violation of natural kinds and artifacts. This is not as clear for nominal kinds (which are by definition artificial constructs). As Keil (1989) claimed, nominal and natural kinds form a continuum (see Fig. 3.2 therein). A violation to a *purely* nominal kind (e.g. circle) would be counterschematic, but a violation to a *partially* nominal (partially natural) kind, such as “uncle” (Keil 1989) would likely be counterintuitive.

Millikan (2009) uses the term *real kinds* to describe real life categories: those with enough distinctive features to be labelled in a non-arbitrary way. Real kinds encompass not only natural kinds (e.g. animal species), but also artifacts (e.g. Gothic cathedrals) and social roles (e.g. doctor). In other words, a real kind is any category with real life meaningfulness. What is the meaning of “meaningful”? In a review of cross-cultural studies, Ross and Tidwell (2010) reported that a person’s *goals* influence how an entity is categorized, and between people, this can differ along cultural lines (e.g. fishermen who fish for *food* versus those who fish for *leisure* will categorize fish differently, see references in Ross and Tidwell 2010; cf. Colby and Cole 1973). Clearly, there is a grey area on the continuum between “intuitive” (in the Boyer and Barrett sense) and culturally derived “habits of thought” (sensu Ross and Tidwell 2010). The precise shape of a person’s conceptual knowledge will vary from person to person. Some people will vary due to their educational standing in life (Meyer and Land 2005; Hampton 2010). Others will vary due to *expertise*. As an aspiring expert gains knowledge-based expertise about a category, then that person might become less focused on the easily visible aspects and more focused on the unseen functional dimensions of the task/stimuli (Ross and Tidwell 2010; cf. Gobet and Simon 1996; Keil 1989; Russell 2011; we return to the topic of expertise later). “Counterintuitive” is perceived when there is a tension between the observed (mundane everyday knowledge) and reports that violate what you know.

A key point is that the supernatural referents of counterintuitive claims are *not perceived directly* (notwithstanding the “fleeting shadow is seen in a clearing,” Boyer 1994, p. 92; also see Boyer 2001, p. 64, Steadman et al. 2009, and Upal 2010). If something was perceived directly and with regularity, then it would cease to be counterintuitive. As Hume (1748/1993) wrote, there “must be a uniform experience *against* every miraculous event, otherwise the event would not merit that appellation” (§90, emphasis added). Human perceptual and cognitive abilities are shaped by our evolutionary history, and all humans share a species-specific pattern (see Laland and Brown 2002). Personal idiosyncrasies aside, there will be certain beliefs that all people should share (McCauley 2000), and this will pertain specifically to the co-observed world—particularly about the *ubiquitously observed* things in the world

that people can discuss when witnessing them together (cf. Atran and Norenzayan 2004, and Clark and Carlson 1981). Regarding the *non-observables*, people are more likely to disagree. This is true for both unseen mundane processes (e.g. plants feeding on water) and the alleged supernatural.

Consider the WITCH concept. It is *defined* by those unseen magic powers. Without them, the witch is merely a woman. Does the witch genuinely have supernatural abilities? It depends on who you ask. You can answer *yes* (if you believe in witches) or *no* (because witches do not exist in the first place). How do we choose between yes and no? We can utilize the tool of “possible worlds” semantics (Lewis 1978) to formulate an additional answer of yes: *yes* (but only in a world where witches *do* exist) (cf. Akman 2009, pp. 405–406, Bloch 2008; Dodd 2011; Guthrie 1993; Hampton 2010, pp. 305–309, Näreaho 2008 and Palmer 1981, pp. 166–173). Hence, a witch-believer and a non-witch believer could both say “yes” for different reasons. To a non-witch-believer, “witch” is an “empty name”: it does not refer to a real life referent (see Reimer 2001 on the “problem of empty names”, and Akman 2009 on “Russell’s problem”). For a witch-believer, the “witch” name is *not* empty. It refers not only to a real person, but to the unseen powers proclaimed in stories of witches. In the mental world of a person who *does* believe in witches, the concept is “witch” is not unusual. Does that mean that it is no longer counterintuitive?

The concept of a witch may *not feel* counterintuitive if—according to an individual’s world view—“witch” is simply regarded as a *real kind* as defined by Millikan (2009; cf. Upal, 2010). In a witch-believing culture, perhaps *witch-believing* does meet the definitional criterion of “maturationally natural” (sensu McCauley 2000) because the concept is embedded in a society where people routinely think about parallel natural and supernatural worlds operating together (Bloch 2008). If a witch is a normal concept, then the “counter-” would be uninstalled from “intuitive” with respect to witches (cf. Bloch 2005; Upal 2010). Is the Nigerian witch a PERSON concept with two violations? Or, is she a WITCH concept with *no* violations? We cannot answer such a question without contemplating the role of context.

5 Concepts in Context

“A context is a psychological construct, a subset of the hearer’s assumptions about the world. It is these assumptions, of course, rather than the actual state of the world, that affect the interpretation of an utterance” (Sperber and Wilson 1996, p. 15).

Context is very important for determining how newly encountered information is understood. Take the sentence “*my typewriter has bad intentions*” (cited by Palmer 1981, p. 49). The claim is anomalous because we know that typewriters do not have the ability to think. Additionally, as Palmer (1981) states, if you “replace *typewriter* by *dog*, *snake*, and *microbe*... Whether the resulting sentence is judged to be anomalous can be determined only by what we know about the intelligence of dogs, snakes, and microbes. The anomaly depends... on knowledge of the world” (p. 50, italics original). In other words, you define strangeness by referencing prior knowledge (Upal et al. 2007; cf. Kintsch 1980). Background knowledge is never static. It can imply different things at different times. Sperber and Wilson (1996) define a *contextual implication* (pp. 107–108) as a situation whereupon an

original context is altered by additional information, which may alter how the referent is understood (cf. Dennett 1995, chapter 14). You might need to keep two premises in your mind simultaneously in order to interpret an utterance (Sperber and Wilson 1996). For example, consider statement (5) from earlier (“a cat has been seen floating on air”). In order to decide whether it is counterintuitive or not, you need at least two premises. If you pair one premise (“the movement of cats is bound by gravity”) with a second premise (“the cat lives on Planet Earth”), then it *is* counterintuitive. Alternately, if you pair that first premise with a third premise (“the cat lives on the International Space Station”), then a floating cat is *not* counterintuitive at all (compare to Gregory and Barrett 2009). In addition to contextual implication, context alters meaning in at least two other ways (Sperber and Wilson 1996): *strengthening* (adding information that supports pre-existing information), and *contradiction* (where new information displaces previous assumptions).

The contextual implication, strengthening, or contradiction to your background knowledge (Sperber and Wilson 1996) comes primarily from your social reference group (Goldberg 2010; Upal 2011a). How does the social referencing differ between science and religion? Religious persons view themselves as inhabiting “truth-seeking communities” (Welker 2006, p. 557) wherein they employ theological explanations of how the world works (ibid., Hume 1748/1993; Guthrie 1980; Masse et al. 2007; McCauley 2000; Pyysiäinen 2003; Park 2005; Vidal and Kleeberg 2007; Whitehouse 1992)—often additionally relying on their own introspection as a source of information (Grim 2006; Whitehouse 2004b) or even using empirical data to justify conclusions about their theologies (as scientists in the Middle Ages did, see Vidal and Kleeberg 2007). Religious claims are, by definition, *not scientific* because they are regarded by their own practitioners as immune (McCauley 2000) from the scientific criterion of falsifiability through observation (Popper 1968). Theological doctrines consist of empirically inscrutable claims embedded within complexes of information that may resemble science in ambition and scope, but which differ crucially from science in how they collect information and update theories (for discussion, see Atkins 2006; Cho 2013; Upal 2010; cf. Dennett 1995, pp. 149–155, Guthrie 1980; Popper 1968; Tweney 2012, and Whitehouse 2004b, Fig. 13.1 therein). However, even the most scientifically rigorous person learns about scientific facts mainly through testimony (from their socially accepted sources), in a manner analogous to how a religious person accepts the testimony of a guru (Audi 2002; cf. Barrett 1999; Hampton 2010; Goldberg 2010, and Upal 2011a). For example, a person only learns that the universe is apparently *isotropic* (same structure in every direction) after reading a prestigious cosmology textbook written by strangers (e.g. Hawley and Holcomb 2005) that makes the claim because other strangers have used a satellite receiver to ascertain the uniformity of microwave radiation is every part of the sky (Hawley and Holcomb 2005). The average person cannot possibly verify this information first hand without access to a microwave radiometer (and the expertise to interpret its output). An atheist and creationist merely disagree on *whose* testimony to accept (cf. Guthrie 1980; also see Shtulman 2013; we return to the topic of testimony later).

In addition to background knowledge, a person also draws inferences based on the real time situation. Theorists in the “externalist” school (see Barrett 2011; Clark 1998; cf. Audi 2002, and Markman and Dietrich 2000) claim that knowledge is not entirely in a person’s head—but that a person’s knowledge base (less concrete than commonly believed) is incomplete without continuous feedback from the environment (which includes cues from one’s own body). In other words, the external environment provides

a scaffold for thinking. Accordingly, semanticists like Akman (2009) have delineated an approach to studying word meanings in a framework that takes into account not only the intrinsic meaning of the word itself, but also the situational context. The hearer needs to make active inferences about the meaning of uttered words based on their selection of *relevant* information from context (Akman 2009; Sperber and Wilson 1996). Hearers “are interested in the meaning of the sentence uttered only insofar as it provides evidence about what the speaker means” (Sperber and Wilson 1996; p. 23). Every communicative situation has a speaker and hearer, their relationship, their consonant (or dissonant) views of the world, and the concepts themselves (Sperber and Wilson 1996, cf. Clark and Carlson 1981).

We have identified three types of context: (1) background semantic knowledge, (2) social reference group, and (3) conversational setting. Each of these should influence how a person develops their world view about counterintuitive concepts. Supernatural stories—a defining feature of religion (Atran and Norenzayan 2004; Pyysiäinen et al. 2003; Steadman et al. 2009)—are memorable, fascinating, and sometimes terrifying (see Boyer 1994, 2001; Jackson 2007), but they are only a small part of everyday religious life (Atran and Norenzayan 2004; Barrett 1999). Religions are more than just a semantic twilight zone (Boyer 2001; Pyysiäinen 2003, 2004): much of religious information is mundane. Moreover, many of the supernatural stories that *do* circulate are “minimally counterintuitive” (Boyer 1994, 2000, 2001). Why?

6 Minimally Counterintuitive Concepts

How do you measure “counterintuitive”? Barrett (2008) has provided a formal methodology. It is a valuable and impressive contribution because it provides a yardstick for *fuzziness* (Upal 2010; cf. Smithson and Verkuilen 2006; also see counter-intuitiveness index used by Johnson et al. 2010). Using Barrett’s (2008) framework, you can quantify counterintuitiveness by counting the number of violations (breaches or transfers) that a concept has. For example, the notation for “invisible dog” is “DOG^p” (the superscript “p” indicating one “physicality” breach to the concept of “dog”), with a counterintuitiveness score of 1. This is called “minimally counterintuitive” (MCI) (Boyer 2001; Barrett 2007b, 2008; Upal 2010, 2011a). If a concept has *no* violations, then it is merely “intuitive” (INT). If a concept has two violations—e.g., “invisible dog that passes through walls”—then it is coded as “DOG^{p+p}” (two physicality breaches). Concepts with two or more violations might be called “maximally counterintuitive” (MAX) concepts (although nobody has proposed a specific criterion for “maximal”, Upal 2010) (some studies refer to it as MXCI, but here we will use MAX). In Barrett’s (2008) scheme, there is no theoretical upper limit to the counterintuitiveness score. As an extreme example, he notates the example of “God” as “HUMAN^{s+p+p+p+b+b+b+m+m+m}”, with a counterintuitiveness score of 10, showing violations to the core concept of human within the domains of physicality, spatiality, biology, and mentality (for reviews on the concepts of God, see Barrett 2007b; Boyer 2001; Dennett 1995; Hick 1983, pp. 5–14; Pyysiäinen 2004, 2005; Shtulman 2010; Slone 2005).

Drawing on our earlier discussion about defining a “concept” (see Medin and Rips 2005; Murphy 2010), we might conclude that “counterintuitiveness” is—in itself—a *relational* concept (i.e. DEVIANCE from known category). Additionally, you can potentially conclude that counterintuitiveness is also a *category* in itself. Categories,

by definition, specify at least one criterion for membership in that category (Murphy 2010). For a concept to be included in the *category* of “counterintuitive”, it needs to have at least one deviation that violates the aforementioned “nomic constraints” (Akman 2009) imposed by the to-be-violated category (cf. Boyer 2000; Barrett 2008). A single concept—such as an invisible dog—will simultaneously be a member of many categories: dog, friend, mammal, living thing, solid object, and—*because of its invisibility*—counterintuitive (Barrett 2008). Yet, category inclusion should not be all-or-nothing. In fuzzy set theory (Smithson and Verkuilen 2006), the assertion is made that different exemplars can have *different degrees* of membership in a category set. For example, a man who is 5'5" tall will have a lesser claim to being included in the category of “tall” than a man who is “6'6”. In other words, both men are “tall”, but one will naturally claim a greater degree of membership in that category (cf. Keil 1989). This viewpoint should also be adopted when we talk about counterintuitiveness as a category. An MCI concept will have less membership in that category than a MAX concept. Barrett's (2008) system allows the grading of concepts regarding *how* counterintuitive they are (cf. Upal et al. 2007). But why should we grade concepts this way?

The gradedness appears to matter very much in cognitive terms. People seem to remember things better if they are *slightly* odd rather than *very* odd. This was known more than 50 years ago:

“We are indifferent to things that are either too remote from our experience or too familiar. A relatively slight variation in a familiar pattern has a unique piquancy. A side show at a fairground that has a two-headed lady on display may well attract more customers than one offering a collection of geological specimens” (Berlyne 1960, p. 21).

Berlyne (1960) studied arousal rather than memory. He described situations where new information caused “conflict” for the perceiver, and this raised arousal and curiosity. In a similar vein, Festinger (1957) described “dissonance” between two simultaneously held beliefs as a cause of discomfort, compelling the cogitators to escape the dissonance by either changing their minds, changing their behaviours, or seeking new information either to support one belief or disconfirm the other. Berlyne (1960) and Festinger (1957) did not use the word “counterintuitive”, but they were essentially claiming that counterintuitiveness grabs attention and creates a need to resolve the issue (cf. Sperber and Wilson 1996; Atran and Norenzayan 2004; Upal 2010). Berlyne (1960) was a pioneer in the study of counterintuitive concepts (e.g. see his empirical study on p. 296), and it was not until many years later that psychologists began to study counterintuitive concepts as it applied to religion. Although the psychology of religion has a long history (Strunk 1971), the study of *memory* for religious material had been mostly ignored in that history (for an early exception, see Morlan 1950/1973). In accordance with the cultural epidemiology approach (Sperber 1985), CSR researchers began to focus on how the properties of the materials itself are conducive to the success of ideas (cf. Whitehouse 2001, 2004a). A frequent observation is that, in established *doctrinal* religions, followers often hold a theological world view that is conceptually far simpler than the official theological writings (see Pyysiäinen 2004; Whitehouse 2004a, b). The implication is that people deviate towards a cognitive optimum. MCI concepts are thought to represent this optimum.

The modern line of research in counterintuitive concepts began with Boyer (1993, 1994, 2001) who drew insights from his own ethnographic field work and emphasized the

importance of counterintuitive concepts in the transmission of religious concepts. In an early experiment, Barrett and Nyhof (2001) showed stories to participants and then asked them to recall the material, finding that counterintuitive concepts are better recalled than regular concepts. Additionally, they found a similar (but less potent) advantage of “bizarre” concepts (i.e. odd but not counterintuitive). Boyer and Ramble (2001) adopted this approach in a cross-cultural study where they set out to investigate whether there were universal patterns in remembering counterintuitive concepts by conducting memory tests in three different cultures (France, Gabon, Tibet). In their first experiment, the participants were exposed to a science fiction story about a museum of items from another planet. The story described 24 items, 12 of which were counterintuitive “breaches” (e.g. floating furniture) and 12 of which were normal (e.g. normal furniture). Furthermore, the items were divided into descriptions of people and descriptions of artefacts. Participants were asked to read the story, and then (after a distraction task), they were asked to recall as many items as possible. The second study was the same except that it tested “transfers” (e.g. people made of metal). In both studies, recall was superior for counterintuitive concepts compared to normal concepts. Also, there was more of a difference between counterintuitive and normal in the “person” category than the “artefact” category. In their third experiment, Boyer and Ramble (2001) decided to compare three different types of concepts: INT concepts (no violations), MCI concepts (one violation, either a breach or transfer), and MAX concepts (with *both* a transfer and a breach). They found that MCI *breaches* were recalled better than MCI transfers, INT concepts, or MAX concepts (this effect was confirmed by Norenzayan et al. 2006). Boyer and Ramble (2001) then replicated their own results in other cultures.

Since 2001, the study of counterintuitive concepts has proliferated (for a historical review, see Upal 2011a). This research has fallen into three main approaches. The first approach is where researchers have scanned the various stories of the world (or else stories created by their own participants) in order to ascertain the incidence of counterintuitive ideas (Barrett et al. 2009; de Cruz 2013; Norenzayan et al. 2006; Tweney et al. 2006; Smith 2009; Upal 2011b; Stubbersfield and Tehrani 2013), generally finding that minimally counterintuitive stories seem to be more optimal than either mundane or maximally counterintuitive stories (judging from the popularity of the stories). The investigators typically did not create new concepts, but examined those created by others. The second research approach (e.g., Banerjee et al. 2013; Barrett 2008; Barrett and Nyhof 2001; Boyer and Ramble 2001; Fondevila et al. 2012; Gregory and Barrett 2009; Hornbeck and Barrett 2013; Johnson et al. 2010; Norenzayan et al. 2006; Slone et al. 2007) is where investigators have focused on the qualities inherent to the concepts themselves—typically involving memory testing for concepts of varying degrees of counterintuitiveness and for the most part confirming the mnemonic advantages of MCI concepts. Here, the investigators usually (but not always) create entirely new concepts to use in their studies. Table 1 is a list of examples of the actual counterintuitive stimuli that have been used in such studies. As shown, not all of them conform to the stipulations of Barrett (2008). In some studies, the participants were given broadly counterintuitive categories (e.g. alien) and asked to fill in the characteristics. Finally, the third research approach (Harmon-Vukić and Slone 2009; Gonce et al. 2006; Upal 2010, 2011a, b; Upal et al. 2007; Harmon-Vukić et al. 2012) is an extension of the second, but with a greater focus on the effects of *context* (see Table 1 for stimuli). Would MCI concepts be optimally recalled, regardless of the contextual background?

The focus on context effects began after Atran and Norenzayan (2004, pp. 721–724) found results that appeared to contradict previous studies (Atran 2006). They had found

that INT concepts were remembered best if the concepts were presented in lists. In an effort to solve this discrepancy, Gonce et al. (2006) investigated the mediating effect of context, finding that context was crucially important for maintaining the optimal memorability of an MCI concept. If they presented the concepts in a list, then INT concepts were recalled best (but not always). If they presented them with an attached descriptive story, then the MCI concept was remembered better than MAX and INT concepts (although the differences between MCI and INT were never significant). They also found that if an INT concept was presented alongside a contradictory context, then it was remembered equally well as an MCI concept. Also, if an MCI concept was presented alongside a contradictory context (i.e. confirming that it was counterintuitive), then it was remembered equally well as an INT concept. Tweney et al. (2006) also investigated context by using the same stimuli used by Gonce et al. (2006), but instead of testing recall, Tweney et al. (2006) asked their participants to choose from the list of concepts and write a complete story (with a beginning, middle, and end) utilizing that concept. They found that participants tended to choose MCI and INT concepts equally often, and tended to avoid using the MAX concepts. Furthermore, they found that there were two different approaches to writing about counterintuitive concepts: *assimilative* (attempting to make the concept normal), or *accommodative* (keeping the concept counterintuitive within a story). In other words, the participants were deciding *how* to put these concepts into context.

Upal et al. (2007) provided the most comprehensive account of context effects. They drew on the ideas of Kintsch (1980), who identified three factors that makes a story compelling to the reader: (1) relevant prior knowledge, (2) expectations that are formed while reading the story (“predictability”), and (3), after finishing the story, the ability to mentally organize the story into a coherent structure (“postdictability”). Upal et al. (2007) provide an explanation that emphasizes the role of computational effort in integrating the new concept into a person’s existing knowledge base (cf. Sperber and Wilson 1996). They claim that MCI concepts are remembered optimally because, even though they are unexpected (low predictability), they can be integrated relatively easily into a person’s existing knowledge base (high postdictability). This is in contrast to MAX concepts which are harder to integrate into the knowledge base because there are more violations. Accordingly, INT concepts are not as memorable because mundane events do not inspire reflection (McCauley 2000).

In the empirical part of their paper, Upal et al. (2007) used three experiments to confirm their model. In their first (a between-subjects design), they had participants rate the postdictability (“ease of justifying”) of a concept after having been asked to assess the meaningfulness of those concepts. Participants had seen either MCI or MAX concepts presented with or without a story attached. The contextual story, in turn, was either supportive (provided a justification for the concept) or contradictory (supported the idea that it was a contradiction). They found that, without a context, participants rated INT concepts as easier to justify than MCI concepts, which were in turn easier to justify than MAX concepts. When the concepts had a contextual story attached, they found that the supportive context generally made the story easier to justify compared to concepts with a contradictory context. When comparing MCI and MAX concepts, they found that the MAX concept with a supportive context was rated as *easier* to justify than an MCI concept with a contradictory context (see Fig. 4 in Upal et al. 2007). Here, the MAX concept had been *rendered less counterintuitive* by the context. In their second experiment, Upal et al. (2007) asked participants to rate the predictability of a particular

concept of occurring in a story. The concepts were either MCI or INT concepts, and the story was intuitive-supportive (supported only INT concepts) or counterintuitive-supportive (supported MCI concepts). They found that participants were conditioned to expect counterintuitiveness during a counterintuitive-supportive story (i.e. in a fairy tale, you expect magic to occur; cf. Skolnick Weisberg and Goodstein 2009). In their final experiment, Upal et al. (2007) used the same recall paradigm (using intuitive-supportive and counterintuitive-supportive stories) to show that higher predictability of a concept made it *less memorable*. This result was complemented by Harmon-Vukić and Slone (2009), who presented the three types of concepts (INT, MCI, MAX) into two types of stories. The first type of story was *integrative* (made the concept follow from what happened in the story) and the second type was *non-integrative* (the concept did not follow). They found that integration made the concepts more memorable, but that presenting the concepts in a story eliminated the differences between the concepts. They speculated that “integration may override the effect of distinctiveness” (ibid., p. 67). Using a similar paradigm, Upal (2011b) found that counterintuitive concepts were more memorable when they had a purpose for the overall story. Similarly, Russell et al. (forthcoming) also found context effects when they asked participants to create stories specifically to accommodate the counterintuitive concepts: the memory difference between INT, MCI, and MAX concepts were diminished. The context-providing procedure of the aforementioned studies (Gonce et al. 2006; Harmon-Vukić and Slone 2009; Upal et al. 2007; Upal 2011b) confirmed the potency of what Sperber and Wilson (1996) called contextual *strengthenings* and contextual *contradictions*.

7 A Question of Expertise

Why does MCI optimality occur? According to our earlier discussion, there must be at least three important background factors: (1) evolved cognitive constraints, (2) properties of the concepts themselves, and (3) cultural context. Recent studies by Upal and colleagues (Gonce et al. 2006; Upal et al. 2007; cf. Harmon-Vukić and Slone 2009; see Upal 2010) have illustrated how *prior context* can change everything. Even a highly bewildering concept can be rendered understandable, or at least partly so. For this reason, we believe that some of the recent work on counterintuitiveness is misguided. Although the taxonomy of Barrett (2008) is admirable, it appears to presume the universality (McCauley 2000) of percepts and concepts which are—in fact—*highly context-dependent*. As proposed earlier, counterintuitiveness is relational. The “minimal” status of an MCI concept exists solely because it is one ‘unit’ of deviation away from a familiar concept. As we have discussed earlier, concepts are not uniformly learned across populations. Different people learn different concepts to varying extents. This differentiation is an issue of *expert learning* (for background, see De Groot and Gobet 1996; Ericsson et al. 2006; Gobet 1998; Gobet et al. 2001, 2011; Gobet and Simon 1996; Russell 2011): a person learns the regularities of a particular domain of knowledge, and this alters that person’s perception of what is counterintuitive and what is not.

Consider the ‘range default’ framework developed by Miller and Perlis (1991, 1997; Miller et al. 1996), where expertise is based on declarative knowledge of categories based on *defaults* (the way that the category is normally understood) and *default denials* (knowledge about when the default situation does not hold). Expressed in plain language (i.e. not using their notation), they define the principles as follows:

- 1) The learner knows the typical features of a category referent (e.g., ‘birds typically fly’).
- 2) The learner knows that the typical features represent a *default*, defining all or most of the exemplars in that category.
- 3) When an exemplar of a category is encountered, the learner anticipates that it will hold the default features (e.g., ‘bird *x* can fly’).
- 4) This expectation constitutes a *defeasible inference*: they learner initially classifies the exemplar according to the default expectations—but this default is subject to rescission if new information shows that the default assumption does not apply (e.g., ‘bird *x* is identified as a penguin, which is known to be flightless’).
- 5) When the exception is recognised, then the ‘default is denied’, but applicable only to the exception. The belief in the truth of the overall trend is not denied (e.g., ‘bird *x* is flightless because it is a penguin, but usually birds do fly’).
- 6) Default denials occur due to knowledge of categories-within-categories, and a knowledge of which attributes define which categories (e.g. the penguin category consists wholly of flightless individuals; most other bird categories consist of flyers).

The above principles were meant to describe ‘expert’ reasoning; but what of the novices? In their more general paper, Miller and Perlis (1997) identify various differences between experts and novices according to their ‘range default’ framework. Consistent with traditional characterizations that experts possess not only more but better-organized knowledge about a topic (Boshuizen and Schmidt 1992; Chi et al. 1981; Chi and Glaser 1985; Lesgold et al. 1988; Patel and Groen 1986), Miller and Perlis (1997) define experts as those who—due to greater breadth of knowledge—are cognizant of a greater number of defaults, more default denials, and have a superior ability to place exceptions in context. The latter ability stems from the fact than an expert has knowledge of more categories-within-categories than does a novice. For example, imagine a novice who has only ever witnessed birds that fly. At this point, the default is too general (‘all birds fly’). After being introduced to a penguin, the novice can now possess two defaults about birds: flyers and flightless. However, the next step is to develop categories whereupon every encountered specimen is put into a comprehensible context (e.g. thinking about different species, their incidence, and the typical characteristics of each). How does this apply to witches? We classify her firstly as a woman (whose default lacks the supernatural abilities of the witch). Then—if you believe in witches—you make a defeasible inference (“she is no ordinary woman”). If you do *not* believe in witches, then you place the witch in the context of fiction (as we did with our talking dog example). Using the range default framework, we can re-classify counterintuitive incidents as a case of *default denial*. Hence, the application of Popper’s (1968) ‘there-is-not’ becomes ‘there-is-usually-not’. In the case of an English-speaking dog, the default denial exists in a fictional context. In the case of a miracle, the default denial is in a context which includes a supernatural agent who allows the miracle to occur. The real issue behind counterintuitiveness is *expectation* (what you would normally expect to happen).

Suppose that a person grew up in Antarctica and had never seen birds other than penguins. In that case, the default is that birds *do not fly* and that person has an expectation that all future birds will be flightless. When a flying bird is finally witnessed, it will be a counterintuitive event (“because birds don’t fly, do they?”). For the Antartical man, the flying bird would be the exception to the rule. Moreover, upon witnessing the non-penguin, the Antartical man is likely to be resistant to changing his overall concept of bird. In general, when encountering a counterintuitive description, people are more likely to try to find ways to justify the occurrence

and preserve as much of their pre-existing beliefs as possible (Upal 2011a; cf. Taber and Lodge 2006; Tweney et al. 2006 on accommodative vs. assimilative approaches). A key point in the issue of counterintuitiveness is that belief is not an *off/on switch* (Pyysiäinen 2003): a person's belief does not change suddenly (cf. Taber and Lodge 2006 and Vosniadou and Brewer 1992). Instead, the accommodation of a newly-encountered counterintuitive concept it is a gradual and complex process whereby the implausible idea becomes a plausible idea only after a process of reflection (e.g. see Carey 2011, about "Quinian bootstrapping").

Let us think more about expertise. Just how far can knowledge about penguins go? Think of the *academic ornithologist* who studies penguins and other birds in a long-term scientific career. For them, a penguin will be far more than a simple category comprising flightlessness, birdhood, etc. For example, in their empirical study relating to penguin *evolution*, Elliot et al. (2013) found that, for penguins, the energetic costs of diving were lower for penguins than for other species that did *both* diving and flying (thick-billed murres and pelagic cormorants), which suggested that flightlessness in penguins evolved because, as birds evolve towards an optimal *swimming* wing shape, they move away from an optimal *flying* wing shape—so for penguins living in nutrient-poor environments (e.g. Antarctica) the best evolutionary path was to move towards a swimming shape (incompatible with flying) which allowed them be agile and efficient underwater foragers yet stay as relatively heavy animals on the ground. For scientists who devote their lives to ornithology, the penguin in this evolutionary scenario refers not to one species but to all of the roughly 20 living species of family *Spheniscidae*, all the extinct species, the typical ecological niches of penguins, probable evolutionary history of each species, their mating habits, social organization, anatomical features, and the list goes on for as long the scientist continues his or her career. Surely this profound level of expertise is light years away from the 'folk' conception of a penguin held by a child or the typical adult non-expert. What can we say about the hunches or intuitions of the penguin specialist? What is counterintuitive for that person?

Thinking back to the five domains of children's knowledge identified by Spelke and Kinzler (2007), we can easily identify cases where default denials (sensu Miller et al. 1996) will apply in those domains (also see Carey 2011). For example, the object representation 'defaults' learned by a child (spatiotemporal principles of cohesion, continuity, and contact) might change radically if the child grows up into a physicist who studies neutrinos (Hawley and Holcomb 2005): tiny elementary particles which that can pass through solid objects. This is an example of how mental expertise challenges what it means for something to be 'counterintuitive' (see Fig. 4 which we discuss later). Because of learning, a physicist will *no longer* find neutrinos to be counterintuitive. Neutrinos cannot be observed with the human eye, but they can be observed through special instruments (but even with instruments, they are extremely difficult to detect). *How* is this cognitively different from accepting the testimony of *other people* about unseen events? The machines that measure neutrinos are—in a sense—giving 'testimony' about the existence of these invisible particles which pass through solid objects. Analogously, a religious person might 'give testimony' about a miracle produced by an invisible God (allegedly also invisible and passing through solid objects). How can you logically distinguish the way these two testimonies are cognitively processed? We will stop our examples here, but there are many other 'default denial' examples that could potentially be found for the domains identified by Spelke and Kinzler (2007)—comprising agency and causation, numeracy, space, and social representation. These predispositions are well-established, but fact that humans are capable of

building up expertise about their world, built upon the knowledge of defaults and default denials (Miller and Perlis 1991, 1997; Miller et al. 1996), carries the possibility that every default can someday be denied (Audi 2002), according to the latest discoveries of science. As discoveries in modern physics and cosmology show (Hawley and Holcomb 2005), our understanding even of the laws of physics are being continuously revised and updated; and physics behaves differently in different parts of the universe. It is true that there are a great many mental representations that all humans will share species-wide (McCauley 2000), but this is true only because we all live in the same *context*: planet Earth.

Research in expertise has a long history (Ericsson et al. 2006) and there are many models that could potentially be applied to the study of counterintuitive concepts. For example, the *template theory* of expertise—originally applied to chess—is actually applicable to all domains of expertise (Gobet 1998; Gobet and Simon 1996; Gobet et al. 2001, 2011; Russell 2011). Going back to Fig. 1 (our WITCH concept), we can differentiate between *core* attributes and *peripheral* attributes. A woman has a genetic endowment that will allow her to potentially give birth to a child. This is an immutable characteristic of womanhood, which excludes her from the category of men (i.e. the concepts of PREGNANCY and MAN are *dissonant* sensu Festinger 1957). In contrast, the higher pitched voice—although highly typical of women—does not always apply (it is defeasible). Template theory is based on *chunks* and *templates* (for background, see Chase and Simon 1973a, b; Gobet and Simon 1996; Gobet 1998; Gobet et al. 2001, 2011). In a chess game, a “chunk” is a configuration of chess pieces observed on a chess board. A newcomer viewing a chess game struggles to recognise patterns on the chessboard: they are all new and seemingly randomly positioned. Here, every individual chess piece can be referred to as a chunk (i.e. a chunk consisting of a single piece). In terms of counterintuitiveness, we might call it the “intuitive” chunk: the single piece is the only “chunk” available to the *untrained* eye. A chess expert *sees it differently*: the same chess board consists of recognizable patterns of pieces. A chunk might consist of three or four—or more—pieces. The expert actually *sees* the larger chunks immediately. Why? It is because the chunks have been observed repeatedly, allowing instant recognition of the commonest arrangements. To explain the chess master’s superior performance in chess, the *chunking theory* (e.g. Chase and Simon 1973a, b) was developed, which postulated that the expert, through lengthy experience, has built up a “large database of chunks” (Gobet 1998, p. 118) in his/her memory. *Template theory* was a refinement of chunking theory, postulating that some chunks are recognized but they are not stored in memory as a singular pattern. Instead, an observed chunk is matched to the closest-fitting template (Gobet 1998). A template is like an averaged-out mental representation: the core characteristics remain the same, even if some of the details vary (Gobet 1998). A chess template might have the same few pieces as its core, but with “movable slots” allowing some variation in the peripheral pieces. If the *core* of the template is violated somehow, then it is no longer the same template.

Let us apply the *template* to the WITCH category. As depicted in Fig. 1, we construe this as a variant of the WOMAN concept. Analogous to that of chess templates, we develop a WOMAN template through repeated exposure throughout the lifetime. The core attributes of womanhood are those without which she could not be classified as a woman (XX chromosomes, childbearing potential, etc.). The peripheral concepts are the “movable slots” that are allowed to vary without disrupting our ability to continue referring to her as a woman. For example, we can vary attributes such as hair colour, height, nationality, and age, and none of those would cancel our ability to call the woman a woman. An exhaustive list of

core woman-concepts would be incomprehensibly long, because the woman concept relates to infinite variables in the environment—and the concept itself subsumed by concepts about the animal kingdom in general (cf. Atran 1998). In other words, there are a large number of superordinate concepts (sensu Palmer 1981): she is not only a woman, but also a human being, a member of genus *Homo*, a primate, a mammal, a chordate, an animal, a biological organism, a carbon-based entity, located on planet Earth, and living in the universe (and the list goes on indefinitely). Hence, we derive a seemingly infinite list of ‘there-is’ and ‘there-is-not’ statements (sensu Popper 1968), and typicality statements (sensu Miller et al. 1996).

A WITCH who “can split her body and soul in two” (Lainé 2007, p. 15) is a complex idea because it has roots in many sources. It refers to cultural beliefs about the supernatural (i.e. that a soul exists) and refers to unseen events based on testimony (more about testimony below). After hearing about this soul-and-body-splitting witch, we generate an ad hoc ‘there-is-not’ statement: “women cannot split their body/souls in two”. We can also extend the proscriptive statement along the hierarchy of concepts (“a human cannot...”, “an animal cannot...”, and so forth). Upon encountering the witch-story, we need to consider the *core concepts* in relation to the *peripheral concepts* (see Fig. 1). Do the body-and-soul-splitting powers violate the core WOMAN concepts (xx-chromosomes, childbearing ability, etc.)? No, they do not. No violation of ‘there-is’ has occurred. If that had occurred, then we have an instance of what Barrett (2008) called *counterschematic*. The key question is whether it is a violation of a ‘there-is-not’ statement (i.e. it is an impossible occurrence), or whether it is simply a report of an *atypical* event. How do you make the distinction? From the point of view of our psychology of expertise argument, *it is simply an issue of typicality*.

This is because impossibility—by definition—is about things that never occur. When a learner hears about an event that was previously thought “impossible”, the event is transformed into “possible” (even if incredibly rare); it is incoherent to posit a concept that is simultaneously possible and impossible. In this article, we construe religious knowledge as simply a particular type of expertise: individuals within a faith have learned about the existence of supernatural entities and they have developed *templates* and *range defaults* to account for these phenomena. The WITCH is therefore linked to the WOMAN template, and the special powers are merely the *moveable slots* on the periphery of the template (she is still a woman, but a special kind of woman). Just as an infant needs time to learn the difference between a MAN concept and a WOMAN concept (Arthur et al. 2009), a person in a religious tradition needs time to learn all of the relevant supernatural concepts in his/her religion. Here, the witch simply becomes a *real kind* (see Guthrie 1993). Another key point, perhaps, is how the information is transmitted to an individual in the first place. Specifically, we should look at the issue of whether the information was gained *directly* or *indirectly*. We address this issue in the next section. First, let us proffer three issues that future studies in counterintuitiveness should *not* ignore:

- (1) The expertise that individuals develop about concepts in their own world (as explained through the explanatory systems prevalent in the surrounding culture), which entails the formation of default knowledge.
- (2) The context where the concept is viewed, including occasions where default denials are expected.

- (3) The source of the default information (experienced directly and indirectly) and how this influences the formation of defaults.

8 Direct Versus Indirect Experience

The third point above (direct versus indirect experience) is particularly interesting line to pursue for defining the nature of counterintuitiveness. The formation of belief is not confined only to an individual’s head (Audi 2002; Goldberg 2010; cf. Clark 1998; cf. Barrett 2011). People rely heavily on the testimony of others to acquire knowledge about the world (Audi 2002; Popper 1968). Testimony can be considered a ‘basic’ source of information because it provides a kind of ‘raw data’ that a person needs to cogitate upon (Audi 2002)—but testimony also carries the burden of additional processing because the hearer of testimony needs to make an assessment about the sanity and honesty of the person giving the testimony (Audi 2002; Goldberg 2010). Can a person who claimed to personally witness a witch’s magical powers be believed? It depends on whether you believed in witches beforehand, and whether you believe the testimony has emerged from a fully sane person with reliable cognitive processes (or else from a group with reliable ways to gather information). An equally important issue in assessing a counterintuitive claim is to reflect on the *lack of testimony* about a particular event (i.e. “if dogs really could speak English, then I would have heard about it now”, see Goldberg 2010). Although a belief is a ‘cognitive achievement’ located solely in the head of an individual, the *origin* of this belief is inextricable from the social and physical context from whence the information came (Goldberg 2010; cf. Clark 1998, and Upal 2011a).

Figure 2 is a diagram of how typicality might be judged according to the relation between direct and indirect sources of knowledge. The horizontal dimension indicates whether the event (e.g. “dog that speaks English”) has been personally witnessed or not. The vertical dimension indicates whether the occurrence has been alleged to have occurred via third party gossip (this gossip may not necessarily be true). This creates four boxes: YY (witnessed + gossip), YN (witnessed, no gossip), NY (not witnessed, gossip), and NN (not witnessed, no gossip). The leftmost two boxes (YY, YN) refer to events regarded as ‘normal’: they are expected to occur because they have been witnessed (whether or not the person has heard stories about these occurrences). The

	Personally witnessed (yes/no)	
Knowledge through gossip (yes/no)	Yes – Yes everyday events	No – Yes fiction, miracles, news about unseen events
	Yes – No everyday events	No – No inconceivable

Fig. 2 Judgment of typicality according to direct information (personal experience) and indirect information (gossip)

bottom right box (NN) refers to events neither described nor witnessed. The upper right box refers to events that have been heard about, but never personally witnessed in real life settings. A “dog that speaks English” would occur in this box. Assuming that witnessed events are more convincing than events heard about second hand, the judged typicality of events would follow this pattern:

$$YY > YN > NY > NN$$

Consider the non-supernatural case of a dog attacking a child. In the YY box, a person might overestimate the actual occurrence of such dog attacks due to the double sources (direct and indirect experience). In the YN box, the typicality might still be regarded as high, but there is only the direct source. In the NY box, the attacks will have been reported but never personally witnessed. Here, the person knows that a dog attack is a possible occurrence, but the lack of personal experience may lead to lower expectations of typicality. In the NN box, there have been no dog attacks witnessed or reported.

Where does a miracle fit into these boxes? An atheist will place a miracle in the NY box (the atheist assumes that miracles do *not* exist, but are spoken about). In contrast, a religious person may choose the YY box (the believer assumes that the miracle has been witnessed by somebody somewhere, and might someday be witnessed personally). This exercise highlights a problem with Fig. 2: *rate of occurrence* is not taken into account. If a religious person places a miracle into the YY box, then it is still considered a rare event. Therefore, we should instead consider the interrelationship between direct and indirect events. Figure 3 is a notional diagram showing the incidence of direct events (x-axis) versus indirect events (y-axis). At the top-right, the incidence of both direct and indirect experience is nearing 100 %, whereupon the event is regarded as highly regular (e.g. that dogs bark, something that is both witnessed personally and frequently heard about). The perception of regularity still exists in the lower-right of the diagram, where the event might be personally witnessed often, without having heard stories about it from other people. On the left of the diagram are events that tend to be heard about but not witnessed. On the bottom left is the conjectured region of ‘counterintuitiveness’. The solid black line indicates the *atheist’s* viewpoint about supernatural events (they do *not* occur, but they are reported by others as relatively rare events; the frequency $0 < y < 50$)

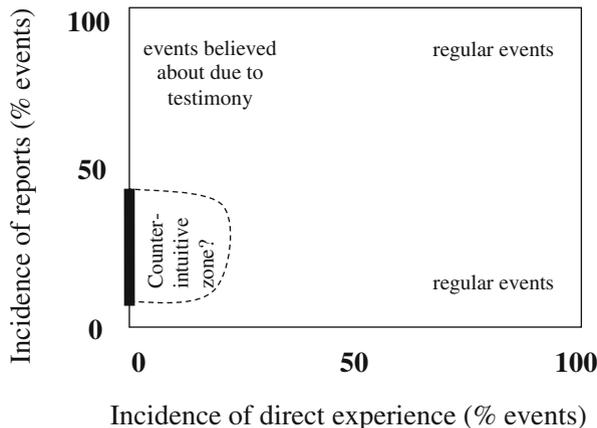


Fig. 3 Counterintuitiveness judgments as a continuous variable according to relative incidence of direct and indirect sources

conjectural). The dotted area illustrates the *religious believer's* perspective about supernatural events: they *are* sometimes witnessed by some people somewhere. Further research is needed in order to quantify exactly where the concepts become counterintuitive, and when such events become frequent enough to be considered merely 'intuitive' (e.g. by using fuzzy set analysis, Smithson and Verkuilen 2006).

9 Folk Psychology vs. Expertise

Think about Earth's roundness. Is it counterintuitive to say that "the Earth is round", because it *looks flat* to uneducated eyes? Research in developmental psychology (Vosniadou and Brewer 1992) has shown that children often initially assume that the world is flat, but eventually change their minds in accordance with lessons from adult culture (but they do not change their minds instantly; there is a gradual process of reconciliation between what adults tell them and the flatness of the landscape they see with their own eyes) (cf. Taber and Lodge 2006). We would argue that the flat earth is *not fully* counterintuitive. Knowledge that the earth is round encourages deeper observation to look for the signs of roundness (e.g. ships disappearing over the horizon). Here-and-now perceptions can be changed simply due to knowledge and expertise. This is why the chess experts discussed earlier can look at a chessboard and automatically perceive various patterns than novices are also 'seeing' but not perceiving (Chase and Simon 1973a, b; De Groot and Gobet 1996; Gobet 1998; Gobet and Simon 1996; cf. Russell 2011). It is true that even simple perceptions can be violated by knowledge ("I see x, but I know that not-x", see Audi 2002), leading to the "counter-perceptual" experiences like visual illusions (Block and Yüker 1992). However, the CSR literature (e.g. Barrett 2008) clearly stipulates that counterintuitiveness is about *new knowledge* contradicting *old knowledge*—not "counter-perceptual" knowledge contradicting perception. When researchers refer to 'folk knowledge', they imply a 'default knowledge' common to everyone (McCauley 2000); but is the word "folk" simply another way of saying *non-expert*? If so, then we need to know where the boundary is. Where does default knowledge end and expert knowledge begin? Is folk knowledge established during childhood? If so, at what ages do learning override folk knowledge? What if the infant is born on the International Space Station, where many of the Earth-bound defaults (nomic constraints) will be different? What of the default knowledge of our evolutionary precursors? Every species in the animal kingdom evolved to perceive conditions on planet Earth, and therefore it could be expected that there will be a number of commonalities in the cognitive mechanisms that animals have developed in order to interpret the regularities in the world (Macphail 1996). Nonetheless, this view underestimates the importance of the species-specific ecological niche in shaping the cognitive architecture of animals, creating vast diversity in how animals use their minds when interacting in the world (e.g., see Barrett 2011; Bolhuis and Giraldeau 2005; Krebs and Davies 1984; Lorenz 1977; Menzel and Fischer 2011). Humans are distinct from other animals in their extreme cognitive reliance on culture (Tomasello 1999). We all share psychological processes which are human-species-specific (Laland and Brown 2002), but *content-wise*, human experience generates incomprehensible diversity (e.g. Hinnells 1984, on religion). For this reason, we believe it is an untenable position to characterize 'intuitive' ontological categories in a manner that suggests that they are innate cognitive structures common to all humans.

However, we are not claiming a complete lack of innate cognitive structures. Figure 4 can help to illustrate our position in a more nuanced manner. The figure illustrates differing implications for counterintuitive as ramifications of changeability and modularity. The boundaries between quadrants should be interpreted as fuzzy and the variables on the axes continuous. The y-axis refers to changeability: the two extremes are “fixed” (the concept and its internal features remain the same throughout life) and “open” (concept is 100 % changeable). The x-axis refers to *modularity*, which is a potentially relevant issue in counterintuitiveness (Salazar 2010). To use the broadest definition, “modularity” refers to the “components, parts, or subsystems of a larger system” (Csete and Doyle 2002, p. 1665) occurring within the anatomy of biological organisms (see Callebaut and Rasskin-Gutman 2005, for a comprehensive exposition of modularity). These subdivisions—the modules—can be structural, processual, or both (Callebaut and Rasskin-Gutman 2005). To provide an unambiguous example, the organs of the human body are modules that are simultaneously structural and processual and—most importantly, *encapsulated* in both ways. The lungs and heart are both structurally discrete and manifest dissimilar processes which benefit the body in complementary ways. There is no benefit to mixing the heart and lungs together. Indeed, modularity is extremely important for the survival of any complex entity (Csete and Doyle 2002)—yet, even more important are what Csete and Doyle (2002) call *protocols*, which are “rules that prescribe allowed interfaces between modules, permitting system functions that could not be achieved by isolated modules” (p. 1666). The heart and lungs may be modular but they don’t function in isolation. In terms of process, the key descriptor of the module is the *encapsulation*—processual independence from other modules and protocols in the system (Csete and Doyle 2002; Callebaut and Rasskin-Gutman 2005) and therefore—returning to Fig. 4—we note that the x-axis refers to the extent of encapsulation of the *concept-as-module*. However, we add the disclaimer that we do not believe in the

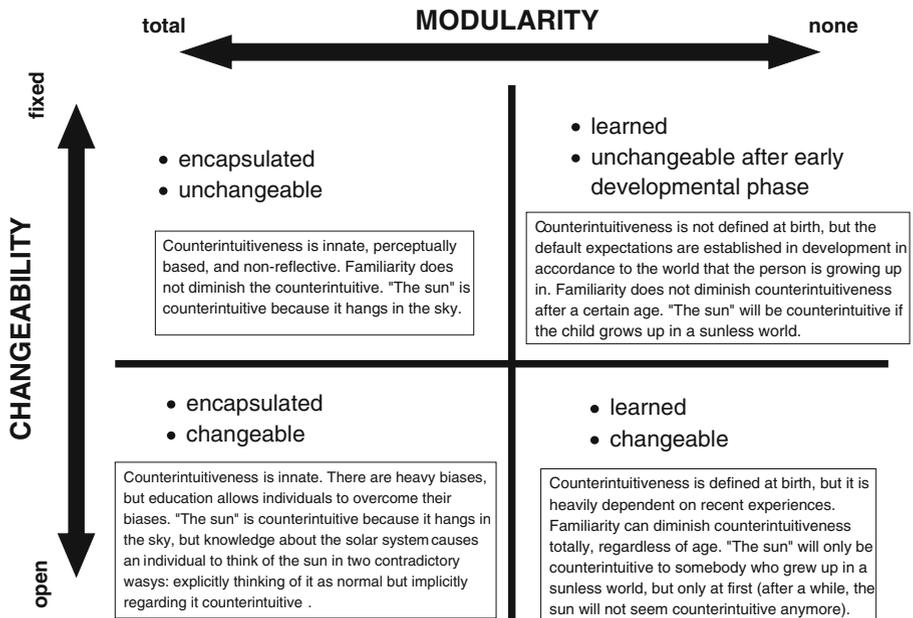


Fig. 4 Counterintuitiveness defined according to changeability and modularity

“concept-as-module”! Consequently, in this paper, we position ourselves somewhere towards the middle of bottom right quadrant, where the perception of counterintuitiveness is subject to change even within a single lifetime.

Why don't we believe in conceptual modularity? The reason is that we regard counterintuition as a highly semantic phenomenon and that the claim that semantic structures are modular is misguided. This is not true of earlier applications of modularity to cognitive processes. Early theorists who spoke of modularity in psychology were referring to *processes* in the mind (see Applebaum 1998), not to the actual *content* that arose out of these processes. While there are highly plausible arguments for the extensiveness of innately-determined processes (Callebaut and Rasskin-Gutman 2005), we cannot see how the concepts themselves can exist as encapsulated “semantic” modules. There are things that can be called core knowledge (Spelke and Kinzler 2007), but these are derived from typical stimuli in the context of Earth-bound life (cf. Newcombe et al. 2009). Language learning, for example, is often cited as having a strong *processual* innateness, but the actual syntactic, lexical and semantic content is built up through exposure to the information and the networks of meaning are built accordingly (Freudenthal et al. 2006, 2007; Gobet et al. 2007; Jones et al. 2000). We believe that the putative universals suggested by Barrett (2008) and others are—in fact—default concepts and percepts which are *defeasible* (cf. Bloch 2005; cf. Cho 2013). Even the ‘nomic constraints’ (Akman 2009)—natural laws as we perceive them—are subject to change as our scientific knowledge about physics grows (see Hawley and Holcomb 2005). That is the whole purpose behind science (Dunbar 1995; Popper 1968): to create knowledge which can be updated through new observations (and find conclusions not obtainable through quick here-and-now perception). Although ‘real kinds’ do exist (Keil 1989), they are subject to revision with additional knowledge (e.g. contextual implications, strengthenings, and contradictions, Sperber and Wilson 1996). Moreover, our ‘default’ knowledge of the world is not *only* the product of first-hand experience, but is derived from testimony given by others (Goldberg 2010). We rarely base our judgments of ‘natural’ or ‘real’ kinds on direct experience alone. With regard to *always-unseen* entities (e.g. a ghost), the problem of subjectivity is compounded—and even for the regularly observed entities (e.g. a cat), there is still subjectivity. Person A and person B may have similar concepts of what a cat can do (it cannot fly, it cannot speak English, it can jump high, it can catch mice), but the similarity exists only insofar as person A and person B have been exposed to congruent information about cats. A zoologist specializing in cats will have a more elaborate conception of a cat than a non-specialist will (e.g., see Turner and Bateson 2000). It is impossible for two people to have an identical conception of a cat, because everyone has a slightly different *experience set* with a concept (zoologists versus non-zoologists being a highly conspicuous example). Facts about cats (like those of penguins and all other animals) may be organized as a kind of template (Gobet and Simon 1996; Gobet 1998; Gobet et al. 2001, 2011): where there are core elements that define a cat (cf. Fig. 1), but where many of the details are allowed to vary. It is irrational to select one single individual's idiosyncratic conception of a cat and claim that it is the only one true universal concept. Yet, it appears that most investigators of counterintuitive concepts are indeed using their own personal concepts to design stimuli that they claim are putatively universal (some examples in Table 1). Although it is difficult to improve on current schemes and research designs (e.g., Barrett 2008, provides a highly useful framework for researchers to follow), it would be more useful for CSR researchers to talk of cognitive units like schemata and concepts

(Anderson 1980) or expertise and templates (Gobet and Simon 1996) than to speak of modules and innate naturalness. We should turn to biological methods (e.g. see the electrophysiological study by Fondevila et al. 2012) to get our answers.

The processes of expertise—our ability to become experts—are products of our evolutionary heritage (see Russell 2011) and in this sense the expertise is genetically driven. In fact, there have been some interesting proposals about how the capacity for different domains of expert intelligence evolved as human-specific cognitive architecture (e.g. Mithen 1996). However, we stay emphatic that evolved predispositions refer to *process*, not *content*. This is evinced in how human beings can form expertise in a bewildering variety of domains (for examples, see Boshuizen and Schmidt 1992; Chi et al. 1981; Chi and Glaser 1985; De Groot and Gobet 1996; Ericsson et al. 2006; Gobet et al. 2011; Lesgold et al. 1988; Patel and Groen 1986; Russell 2011; Shalin et al. 1997). Such diversity of human brilliance would be simply impossible if our conceptual knowledge were genetically cloistered.

10 Conclusion

Clearly, the boundary whereupon *folk* becomes *non-folk* has not been defined specifically enough. At what stage does a zoology student studying penguins or cats transform into *non-folk*? Moreover, a person's acceptance of counterintuitive ideas is a much subtler and ratiocinative process than is normally assumed (Pyysiäinen 2003). Our knowledge of the world is an incomplete similitude of the way it really is, aided by the ongoing process of revision used in science (Dunbar 1995; Popper 1968; cf. Pyysiäinen 2003; cf. Day 2007); and also because we are constrained by our evolved cognitive processes in the way we process and prioritize information (Bolhuis and Giraldeau 2005; Lorenz 1977; Menzel and Fischer 2011; Upal 2011a) and must rely on testimony to get much of our knowledge (Audi 2002; Goldberg 2010). Perhaps we should abandon the idea that counterintuitiveness is based on the 'nomic constraints' (Akman 2009). Instead, we can regard them as based on 'conventional' constraints (Akman 2009): "truths" that are determined through the collective opinion of a social group (Upal 2011a), no matter how the "truth" was originally established (whether by science or religion). Upal (2010) proffered a very useful distinction between the *concept view* and *context view* of counterintuitiveness. Although there is great research value in a concept-based approach (as reviewed above, the effects are real, and need explanation), it would be naïve to claim that a focus on *concepts in isolation* can tell us everything about counterintuitiveness. The weight of psychological research is more supportive of the context view (background knowledge determines what is counterintuitive). Drawing from the above discussion, we define "counterintuitive" this way:

"Counterintuitive" is a subjective assessment of a single event (experienced directly or indirectly), where the evaluator has judged that category A is applicable—yet in that instance at least one of the properties of category A are *incompatible* with previously known exemplars of category A (experienced directly or indirectly)—but the evaluator chooses to continue applying category A in its previously known form.

In the situation above, the evaluation may be temporary or permanent. In the permanent case, the evaluator may choose to accept the contradiction without further examination (e.g. as a miracle). In the temporary case, after some duration of time—perhaps a millisecond, a

minute, or a decade—the evaluator might choose to reclassify the event into something mundane. As Näreaho (2008) pointed out, some previous definitions of counterintuitiveness appeared to deliberately leave open the possibility that the referent—the supernatural entity itself—actually exists (and moreover he opines that this creates a clash with the stated naturalistic agenda of CSR). In our definition and discussion above, we close off the possibility that these entities really exist (and if they do, they are simply a description of highly atypical occurrences). In other words, if we hear that someone believes in x —yet there is no way for us to observe the x —then, as scientists, we can only study the beliefs about x —not the x itself (Steadman et al. 2009).

Our aim in this review was to mount a strong challenge to innatist assumptions made in the recent literature. Our conclusions are congruent with Bloch (2005) who wrote: “if some of the representations are so familiar as to be just like intuitive beliefs, then the special catchiness attributed to the stimulation of the counter-intuitive will be nullified” (p. 115). Rebuttals to his statement (e.g. Lanman 2007) claim that concepts can become familiar yet still remain catchy and distinctive—but we cannot see how this claim gels with the ubiquitous human custom of becoming an *expert* where formerly unhabituated concepts become incorporated into a person’s knowledge and skill set. We agree with Hume (1748/1993) that a ‘routine miracle’ is no longer a miracle. In our view, a ‘counterintuitive’ event is a tailor-made surprise, a confounding of default expectation that is unique to the individual—even if it parallels the surprise of many other people with similar life experiences. If the conditions of the world change—then so would our defaults.

Acknowledgments (YIR): I thank my colleagues in the Explaining Religion (EXREL) project, especially Harvey Whitehouse for organising the whole project, the other research fellows, PhD students, other project partners, and administrative staff. Thank you to Justin Barrett for enthusiastically disagreeing with me in a seminar. I also thank Afzal Upal, Nicolas Baumard, Jennifer Dolinsky, and Harvey Whitehouse for comments on a much earlier and shorter draft, and to Dirk Semmann for allowing me time to work on this manuscript in Germany. Y.I.R. was partly funded by the German Initiative of Excellence of the German Science Foundation (DFG) and by the EC-FP6 EXREL project grant 43225.

References

- Akman, V. 2009. Situated semantics. In *The Cambridge Handbook of Situated Cognition*, eds. P. Robbins, and M. Aydede, 401–418. Cambridge: Cambridge University Press.
- Anderson, J.R. 1980. Concepts, propositions, and schemata: What are the cognitive units? *Nebraska Symposium on Motivation* 28: 121–162.
- Applebaum, I. 1998. Modularity. In *A Companion to Cognitive Science*, ed. W. Bechtel and G. Graham, 625–635. Oxford: Blackwell.
- Arthur, A.E., R.S. Bigler, and D.N. Ruble. 2009. An experimental test of the effects of gender constancy on sex typing. *Journal of Experimental Child Psychology* 104: 427–446.
- Atkins, P. 2006. Atheism and science. In *The Oxford Handbook of Religion and Science*, eds. P. Clayton and Z. Simpson, 124–136. Oxford: Oxford University Press.
- Atran, S. 1998. Folkbiology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21: 547–649.
- Atran, S. 2006. The scientific landscape of religion: Evolution, culture, and cognition. In *The Oxford Handbook of Religion and Science*, eds. P. Clayton and Z. Simpson, 407–429. Oxford: Oxford University Press.
- Atran, S., and A. Norenzayan. 2004. Religion’s evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences* 27: 713–770.
- Audi, R. 2002. The sources of knowledge. In *The Oxford Handbook of Epistemology*, ed. P.K. Moser, 71–94. Oxford: Oxford University Press.

- Banerjee, K., O.S. Haque, and E.S. Spelke. 2013. Melting lizards and crying mailboxes: Children's preferential recall of minimally counterintuitive concepts. *Cognitive Science*. doi:10.1111/cogs.12037.
- Barrett, J.L. 1999. Theological correctness: Cognitive constraint and the study of religion. *Journal for the Scientific Study of Religion* 11: 325–339.
- Barrett, J.L. 2007a. Cognitive science of religion: What is it and why is it? *Religion Compass* 1: 1–19.
- Barrett, J.L. 2007b. Gods. In *Religion, Anthropology, and Cognitive Science*, eds. H. Whitehouse and J. Laidlaw, 179–210. Durham, NC: Carolina Academic Press.
- Barrett, J.L. 2008. Coding and quantifying counterintuitiveness in religious concepts: Theoretical and methodological reflections. *Method & Theory in the Study of Religion* 20: 308–338.
- Barrett, J.L., and F.C. Keil. 1996. Conceptualizing a non-natural entity: Anthropomorphism in God concepts. *Cognitive Psychology* 31: 219–247.
- Barrett, J.L., and M.A. Nyhof. 2001. Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *Journal of Cognition and Culture* 1: 69–100.
- Barrett, J.L., and R.A. Richert. 2003. Anthropomorphism or preparedness? Exploring children's God concepts. *Review of Religious Research* 44: 300–312.
- Barrett, J.L., E. Reed Burdett, and T. Porter. 2009. Counterintuitiveness in folktales: Finding the cognitive optimum. *Journal of Cognition and Culture* 9: 271–287.
- Barrett, L. 2011. *Beyond the Brain: How Body and Environment Shape Animal and Human Minds*. Princeton, NJ: Princeton University Press.
- Barsalou, L.W. 2009. Situating concepts. In *The Cambridge Handbook of Situated Cognition*, eds. P. Robbins and M. Aydede, 236–263. Cambridge: Cambridge University Press.
- Berger, L. 2013. A revised view of the judicial hunch. *Legal Communication & Rhetoric: JALWD* 10.
- Berlyne, D.E. 1960. *Conflict, Arousal, and Curiosity*. New York: McGraw Hill.
- Betsch, T. 2008. The nature of intuition and its neglect in research on judgment and decision making. In *Intuition in Judgment and Decision Making*, eds. H. Plessner, C. Betsch and T. Betsch, 3–22. New York: Lawrence Erlbaum.
- Bloch, M. 2005. Are religious beliefs counter-intuitive? In *Essays on Cultural Transmission*, ed. M. Bloch, 103–121. Cambridge: Cambridge University Press.
- Bloch, M. 2008. Why religion is nothing special but is central. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 2055–2061.
- Block, J.R., and H. Yuker. 1992. *Can You Believe Your Eyes?* New York: Brunner/Mazel.
- Bolhuis, J.J., and L.-A. Giraldeau, eds. 2005. *The Behaviour of Animals: Mechanisms, Function, and Evolution*. Oxford: Blackwell.
- Boshuizen, H.P.A., and H.G. Schmidt. 1992. On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science* 16: 153–184.
- Boyer, P. 1993. Pseudo-natural kinds. In *Cognitive Aspects of Religious Symbolism*, ed. P. Boyer, 121–142. Cambridge: Cambridge University Press.
- Boyer, P. 1994. *The Naturalness of Religious Ideas: A Cognitive Theory of Religion*. Berkeley: University of California Press.
- 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. 2001. *Religion Explained: The evolutionary origins of religious thought*. New York: Basic Books.
- 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.
- Callebaut, W., and D. Rasskin-Gutman, eds. (2005). *Modularity: Understanding the Development and Evolution of Natural Complex Systems*. Cambridge, MA: MIT.
- Carey, S. 2011. Précis of *The Origin of Concepts*. *Behavioral and Brain Sciences* 34: 113–167.
- Chase, W.G., and H.A. Simon. 1973a. Perception in chess. *Cognitive Psychology* 4: 55–81.
- Chase, W.G., and H.A. Simon. 1973b. The mind's eye in chess. In *Visual Information Processing*, ed. W.G. Chase, 215–281. New York: Academic Press.
- Chassy, P., and F. Gobet. 2011. A hypothesis about the biological basis of expert intuition. *Review of General Psychology* 15: 198–212.
- Chi, M.T.H., P.J. Feltovich, and R. Glaser. 1981. Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5: 121–152.
- Chi, M.T.H., and R. Glaser. 1985. Problem solving ability. *Human Abilities: An Information-Processing Approach*, ed. R.J. Sternberg, 227–250. New York: Freeman.
- Cho, F. 2013. Unnatural comparisons: Commentary on Robert McCauley's why religion is natural and science is not. *Religion, Brain, and Behavior* 3: 119–182.
- Clark, A. 1998. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT.

- Clark, H.H., and T.B. Carlson. 1981. Context for comprehension. In *Attention and Performance, Vol. IX*, eds. J. Long and A. Baddeley, 313–330. Hillsdale, NJ: Erlbaum.
- Cohen, E., J.A. Lanman, H. Whitehouse, and R.N. McCauley. 2008. Common criticisms of the cognitive science of religion—answered. *Bulletin of the Council of Societies for the Study of Religion* 37: 112–115.
- Colby, B., and M. Cole. 1973. Culture, memory, and narrative. In *Modes of Thought: Essays on Thinking in Western and Non-Western Societies*, eds. R. Horton and R. Finnegan, 63–91. London: Faber & Faber.
- Csete, M.E., and J.C. Doyle. 2002. Reverse engineering of biological complexity. *Science* 295: 1664–1669.
- Cullen, B. 2000. *Contagious Ideas: On Evolution, Culture, Archaeology, and Cultural Virus Theory*. Oxford: Oxbow Books.
- Day, M. 2007. Let's be realistic: Evolutionary complexity, epistemic probabilism, and the cognitive science of religion. *Harvard Theological Review* 100: 47–64.
- De Cruz, H. 2013. Religious concepts as structured imagination. *International Journal for the Psychology of Religion* 23: 63–74.
- De Groot, A., and F. Gobet. 1996. *Perception and Memory in Chess. Heuristics of the Professional Eye*. Assen, NL: Van Gorcum.
- Dennett, D. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- Dodd, D. 2011. Quasi-miracles, typicality, and counterfactuals. *Synthese* 179: 351–360.
- Dunbar, R. 1995. *The Trouble with Science*. Cambridge, MA: Harvard University Press.
- Elliot, K.H., R.E. Ricklefs, A.J. Gaston, S.A. Hatch, J.R. Speakman, and G.K. Davoren. 2013. High flight costs, but low dive costs, in auks support the biochemical hypothesis for flightlessness in penguins. *Proceedings of the National Academy of the United States of America* 110: 9380–9384.
- Ericsson, K.A., N. Charness, P.J. Feltovich, and R.R. Hoffman, eds. 2006. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press.
- Feldman, J. 2003. The simplicity principle in human concept learning. *Current Directions in Psychological Science* 12: 227–239.
- Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.
- Fondevila, S., M. Martin-Loeches, L. Jiménez-Ortega, P. Casado, A. Sel, A. Fernández-Hernández, and W. Sommer. 2012. The sacred and the absurd – an electrophysiological study of counterintuitive ideas (that sentence level). *Social Neuroscience* 7: 445–457.
- Freudenthal, D., J.M. Pine, and F. Gobet. 2006. Modelling the development of children's use of optional infinitives in English and Dutch using MOSAIC. *Cognitive Science* 30: 277–310.
- Freudenthal, D., J.M. Pine, J. Aguado-Orea, and F. Gobet. 2007. Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science* 31: 311–341.
- Ganea, P.A., K. Shutts, E. Spelke, and J.S. de Loache. 2007. Thinking of things unseen: Infants' use of language to update mental representations. *Psychological Science* 18: 734–739.
- Gobet, F. 1998. Expert memory: A comparison of four theories. *Cognition* 66: 115–152.
- Gobet, F., and P. Chassy. 2009. Expertise and intuition: A tale of three theories. *Minds and Machines* 19: 151–180.
- Gobet, F., P. Chassy and M. Bilalić. 2011. *Foundations of Cognitive Psychology*. McGraw-Hill.
- Gobet, F., P.C.R. Lane, S. Croker, P.C.-H. Cheng, G. Jones, I. Oliver, and J.M. Pine. 2001. Chunking mechanisms in human learning. *Trends in Cognitive Sciences* 5: 236–243.
- Gobet, F., J.M. Pine, and D. Freudenthal. 2007. Towards a unified model of language acquisition. *Proceedings of the European Cognitive Science Conference 2007*, eds. S. Vosniadou, D. Kayser and A. Protopapas, 602–607. New York: Taylor & Francis.
- Gobet, F., and H.A. Simon. 1996. Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology* 31: 1–40.
- Gobet, F., and H.A. Simon. 1998. Expert chess memory: Revisiting the chunking hypothesis. *Memory* 6: 225–255.
- Goldberg, S.C. 2010. *Relying on Others: An Essay in Epistemology*. Oxford: Oxford University Press.
- Gonce, L.O., M.A. Upal, D.J. Slone, and R.D. Tweney. 2006. Role of context in the recall of counterintuitive concepts. *Journal of Cognition and Culture* 6: 521–547.
- Gorsuch, R.L. 1988. Psychology of religion. *Annual Review of Psychology* 39: 201–221.
- Gregory, J.P., and J.L. Barrett. 2009. Epistemology and counterintuitiveness: Role and relationship in epidemiology of cultural representations. *Journal of Cognition and Culture* 9: 289–314.
- Grim, J. 2006. Indigenous lifeways and knowing the world. In *The Oxford Handbook of Religion and Science*, eds. P. Clayton and Z. Simpson, 87–107. Oxford: Oxford University Press.
- Guthrie, S.E. 1980. A cognitive theory of religion. *Current Anthropology* 21: 181–203.
- Guthrie, S.E. 1993. *Faces in the Clouds: A New Theory of Religion*. Oxford: Oxford University Press.
- Guthrie, S.E. 2007. Anthropological theories of religion. In *The Cambridge Companion to Atheism*, ed. M. Martin, 283–299. Cambridge: Cambridge University Press.

- Hampton, J.A. 1999. Concepts. In *The MIT Encyclopedia of the Cognitive Sciences*, eds. R.A. Wilson and F.C. Keil, 176–179. Cambridge, MA: MIT.
- Hampton, J.A. 2010. Concepts in human adults. In *The Making of Human Concepts* eds. D. Mareschal, P.C. Quinn and S.E.G. Lea, 293–311. Oxford: Oxford University Press.
- Harmon-Vukić, M.E., and D.J. Slone. 2009. The effect of integration on recall of counterintuitive stories. *Journal of Cognition and Culture* 9: 57–68.
- Harmon-Vukić, M.E., M.A. Upal, and K.J. Sheehan. 2012. Understanding the memory advantage of counterintuitive concepts. *Religion, Brain & Behaviour* 2: 121–139.
- Hawley, J.F., and K.A. Holcomb. 2005. *Foundations of Modern Cosmology*. Oxford: Oxford University Press.
- Hick, J.H. 1983. *Philosophy of Religion, 3rd edition*. Englewood Cliffs, NJ: Prentice Hall.
- Hinnells, J.R., ed. 1984. *The Penguin Dictionary of Religions*. London: Penguin.
- Hornbeck, R.G., and J.L. Barrett. 2013. Refining and testing “counterintuitiveness” in virtual reality: Cross-cultural evidence for recall of counterintuitive representations. *International Journal for the Psychology of Religion* 23: 15–28.
- Hume, D. 1748. *An Enquiry Concerning Human Understanding*. Reprinted in *An Enquiry Concerning Human Understanding, 2nd edition*, 1993, ed. E. Steinberg. Indianapolis/Cambridge: Hackett Publishing Company.
- Jackson, B. 2007. Jonathan Edward goes to hell (house): Fear appeals in American evangelism. *Rhetoric Review* 26: 42–59.
- Johnson, C.V.M., S.W. Kelly, and P. Bishop. 2010. Measuring the mnemonic advantage of counter-intuitive and counter-schematic concepts. *Journal of Cognition and Culture* 10: 109–121.
- Jones, G., F. Gobet, and J.M., Pine 2000. A process model of children’s early verb use. In *Proceedings of the Twenty Second Annual Meeting of the Cognitive Science Society*, eds. L.R. Gleitman and A.K. Joshi, 723–728. Mahwah, NJ: Lawrence Erlbaum.
- Keil, F. 1989. *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT.
- Kintsch, W. 1980. Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics* 9: 87–98.
- Koltko-Rivera, M.E. 2004. The psychology of worldviews. *Review of General Psychology* 8: 3–58.
- Krebs, J.R., and N.B. Davies, eds. 1984. *Behavioural Ecology: An Evolutionary Approach, 2nd edition*. Oxford: Blackwell.
- Laidlaw, J. 2007. A well-disposed social anthropologist’s problems with the ‘cognitive science of religion’. In *Religion, Anthropology, and Cognitive Science*, eds. H. Whitehouse and J. Laidlaw, 211–246. Durham, NC: Carolina Academic Press.
- Lainé, D. 2007. *African Gods: Contemporary Rituals and Beliefs*. Paris: Flammarion.
- Laland, K.N., and G.R. Brown. 2002. *Sense and Nonsense: Evolutionary Perspectives on Human Behaviour*. Oxford: Oxford University Press.
- Lanman, J. 2007. Review of Essays on Cultural Transmission by Maurice Bloch. *The Journal of Biosocial Science* 39: 633–634.
- Legare, C., A. Shtulman, J. Bering, and H. Whitehouse. 2009. The cognitive science of religion. In *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society*, eds. N. Taatgen and H. van Rijn, 2081–2082. Amsterdam: Cognitive Science Society.
- Lesgold, A., H. Rubinson, P. Feltovich, R. Glaser, D. Klopfer and Y. Wang. 1988. Expertise in a complex skill: Diagnosing X-ray pictures. In *The Nature of Expertise*, eds. M.T.H. Chi, R. Glaser and R. Farr, 311–342. Hillsdale, NJ: Erlbaum.
- Lewis, D. 1978. Truth in fiction. *American Philosophical Quarterly* 15: 37–46.
- Lorenz, K. 1977. *Behind the Mirror*. London: Methuen.
- Macphail, E.M. 1996. Cognitive function in mammals: The evolutionary perspective. *Cognitive Brain Research* 3: 279–290.
- Malt, B.C., E. Ameel, S. Gennari, M. Imai, N. Saji, and A. Majid. 2011. Do words reveal concepts? In *Expanding the Space of Cognitive Science: Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, eds. L. Carlson, C. Hoelscher, and T.F. Shipley, 519–524. Austin, TX: Cognitive Science Society.
- Markman, A.B., and E. Dietrich. 2000. Extending the classical view of representation. *Trends in Cognitive Science* 4: 470–475.
- Masse, W.B., E.W. Barber, L. Piccardi, and P.T. Barber. 2007. Exploring the nature of myth and its role in science. *Geological Society of London, Special Publications* 273: 9–28.
- McCausley, R.N. 2000. The naturalness of religion and the unnaturalness of science, In *Explanation and Cognition*, eds. F. Keil and R. Wilson, 61–85. Cambridge, MA: MIT.

- Medin, D.L., and L.J. Rips. 2005. Concepts and categories: Memory, meaning, and metaphysics. In *The Cambridge Handbook of Thinking and Reasoning*, eds. K.J. Holyoak and R.G. Morrison, 37–72. Cambridge: Cambridge University Press.
- Menzel, R., and J. Fischer, eds. 2011. *Animal Thinking: Contemporary Issues in Comparative Cognition*. Cambridge, MA: MIT.
- Meyer, J.H.F., and M. Land. 2005. Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education* 49: 373–388.
- Miller, M., and D. Perlis. 1991. Typicality constants and range defaults: Some pros and cons of a cognitive model of default reasoning. *Lecture Notes in Computer Science* 542: 560–569.
- Miller, M., and D. Perlis. 1997. Toward automatic expert reasoning and expert novice communication. In *Expertise in Context: Human and Machine*, eds. P.J. Feltovich, K.M. Ford and R.R. Hoffman, 405–416. London/Menlo Park, CA/Cambridge, MA: AAAI Press/MIT.
- Miller, M., D. Perlis, and K. Purang. 1996. Defaults denied. *Technical report CS-TR-3680*. University of Maryland Computer Science Department.
- Millikan, R. 2009. Embedded rationality. In *The Cambridge Handbook of Situated Cognition*, eds. P. Robbins and M. Aydede, 171–181. Cambridge: Cambridge University Press.
- Mithen, S. 1996. *The Prehistory of the Mind: The Cognitive Origins of Art, Religion, and Science*. London: Thames and Hudson.
- Morlan, G.K. 1950. An experiment on the recall of religious material. *Religion in Life*, 19, 589–594. Reprinted in *Psychology & Religion*, ed. L.B. Brown, 1973. Aylesbury, Buckinghamshire: Hazel, Watson & Viney.
- Murphy, G.L. 2010. What are categories and concepts? In *The Making of Human Concepts*, eds. D. Mareschal, P.C. Quinn and S.E.G. Lea, 11–28. Oxford: Oxford University Press.
- Näreaho, L. 2008. The cognitive science of religion: Philosophical observations. *Religious Studies* 44: 83–98.
- Newcombe, N.S., K.R. Ratliff, W.L. Shallcross, and A. Twyman. 2009. Is cognitive modularity necessary in evolutionary account of development? In *Cognitive Biology: Evolutionary and Developmental Perspectives on Mind, Brain, and Behaviour*, eds. L. Tommasi, M.A. Peterson, and L. Nadel, 105–126. Cambridge, MA: MIT.
- Norenzayan, A., S. Attran, J. Faulkner, and M. Schaller. 2006. Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science* 30: 531–553.
- Nynäs, P. 2008. Counter-intuitive religious representations from the perspective of early intersubjective development and complex representational constellations. A methodological reflection. *Archives for the Psychology of Religion* 30: 37–55.
- Palmer, F.R. 1981. *Semantics, 2nd edition*. Cambridge: Cambridge University Press.
- Park, C.L. 2005. Religion and meaning. In *Handbook of the Psychology of Religion and Spirituality*, eds. R.F. Paloutzian and C.L. Park, 295–314. New York: The Guildford Press.
- Patel, V.L., and G.J. Groen. 1986. Knowledge based solution strategies in medical reasoning. *Cognitive Science* 10: 91–116.
- Pawlikowski, J. 2007. The history of thinking about miracles in the west. *Southern Medical Journal* 100: 1229–1235.
- Popper, K.R. 1968. *The Logic of Scientific Discovery, revised edition*. New York: Harper & Row.
- Pyysiäinen, I. 2002. Religion and the counter-intuitive. In *Current Approaches in the Cognitive Science of Religion*, eds. I. Pyysiäinen and V. Anttonen, 110–132. London: Continuum.
- Pyysiäinen, I. 2003. True fiction: Philosophy and psychology of religious belief. *Philosophical Psychology* 16: 109–125.
- Pyysiäinen, I. 2004. *Magic, Miracles, and Religion: A Scientist's Perspective*. Walnut Creek, CA: Alta Mira.
- Pyysiäinen, I. 2005. God: A brief history with a cognitive explanation of the concept. *Temenos: Nordic Journal of Comparative Religion* 41: 77–128.
- Pyysiäinen, I. 2013. Cognitive science of religion: State-of-the-art. *Journal of the Cognitive Science of Religion* 1: 5–28.
- Pyysiäinen, I., M. Lindeman, and T. Honkela. 2003. Counterintuitiveness as the hallmark of religion. *Religion* 33: 341–355.
- Reimer, M. 2001. The problem of empty names. *Australasian Journal of Philosophy* 79: 491–506.
- Ross, N. and M., Tidwell. 2010. Concepts and culture. In *The Making of Human Concepts*, eds. D. Mareschal, P.C. Quinn and S.E.G. Lea, 131–148. Oxford: Oxford University Press.
- Russell, Y.I. 2011. Prehistoric stone tools, chess expertise, and cognitive evolution: An experiment about recognizing features in flint debitage. *Journal of Evolutionary Psychology* 9: 249–269.

- Russell, Y.I., F. Gobet and H. Whitehouse, forthcoming. Recall of 'counterintuitive' concepts: The effect of pre-training, presentation, and normality of concepts.
- Salazar, C. 2010. Anthropology and the cognitive science of religion: A critical reassessment. *Religion and Society: Advances in Research* 1: 44–56.
- Saler, B. 2010. Theory and criticism: The cognitive science of religion. *Method & Theory in the Study of Religion* 22: 330–339.
- Shalin, V.L., N.D. Geddes, D. Bertram, M.A. Szczepekowski, and D. Dubois. 1997. Expertise in dynamic, physical task domains. In *Expertise in Context: Human and Machine*, eds. P.J. Feltovich, K.M. Ford and R.R. Hoffman, 195–217. London/Menlo Park, CA/Cambridge, MA: AAAI Press/MIT.
- Shtulman, A. 2010. Theories of God: explanatory coherence in a non-scientific domain. In *Cognition in Flux: Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, eds. S. Ohlsson and R. Catrambone, R., 1295–1300. Hillsdale, NJ: Erlbaum.
- Shtulman, A. 2013. Epistemic similarities between students' scientific and supernatural beliefs. *Journal of Educational Psychology* 105: 199–212.
- Skolnick Weisberg, D., and J. Goodstein. 2009. What belongs in a fictional world? *Journal of Cognition and Culture* 9: 69–78.
- Slone, D.J. 2005. Why religions develop free-will problems. In *Mind and Religion: Psychological and Cognitive Foundations of Religiosity*, eds. H. Whitehouse and R.N. McCauley, 187–206. Walnut Creek, CA: Alta Mira.
- Slone, D.J., L. Gonce, M.A. Upal, K. Edwards, and R. Tweney. 2007. Imagery effects on recall of counterintuitive concepts. *Journal of Cognition and Culture* 7: 355–367.
- Smith, A. 2009. An exploration of counter-intuitive conceptual structures in organizational stories. *Journal of Sport Management* 23: 483–510.
- Smithson, M. and J. Verkuilen. 2006. *Fuzzy set theory: Applications in the Social Sciences*. Thousand Oaks, CA: Sage Publications.
- Sommerfeld, R.D., H.-J. Krambeck, D. Semmann, and M. Milinski. 2007. Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences of the United States of America* 104: 17435–17440.
- Spelke, E.S., and K.D. Kinzler. 2007. Core knowledge. *Developmental Science* 10: 89–96.
- Sperber, D. 1985. Anthropology and psychology: Towards an epidemiology of representations. *Man* 20: 73–89.
- Sperber, D. and D. Wilson. 1996. *Relevance: Communication & Cognition*. Oxford: Blackwell.
- Steadman, LB, C.T. Palmer, and R.M. Ellsworth. 2009. Toward a testable definition of religious behaviour. In *The Biology of Religious Behaviour*, ed. J.R. Feierman, 20–35. Santa Barbara, CA: Praeger.
- Strunk, O., Jr., ed. 1971. *Psychology of Religion: Historical and Interpretive Readings*. Nashville/New York: Abingdon Press.
- Stubbersfield, J., and J. Tehrani. 2013. Expect the unexpected? Testing for minimally counterintuitive (MCI) bias in the transmission of contemporary legends: A computational phylogenetic approach. *Social Science Computer Review* 31: 90–102.
- Taber, C.S., and M. Lodge. 2006. Motivated scepticism in the evaluation of political beliefs. *American Journal of Political Science* 50: 755–769.
- Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Turner, D.C. and P. Bateson, eds. 2000. *The Domestic Cat: The Biology of its Behaviour, 2nd edition*. Cambridge: Cambridge University Press.
- Tweney, R D. 2012. Heuristics and the counterintuitive in science and religion. In *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution*, eds. K.S. Rosengren, S.K. Brem, E.M.Evans and G.M. Sinatra, 270–283. Oxford: Oxford University Press.
- Tweney, R.D., M.A. Upal, L.O. Gonce, D.J. Slone, and K. Edwards. 2006. The creative structuring of counter-intuitive worlds. *Journal of Cognition and Culture* 6: 483–498.
- Upal, M.A. 2010. An alternative account of the minimal counterintuitiveness effect. *Cognitive Systems Research* 11: 194–203.
- Upal, M.A. 2011a. From individual to social counterintuitiveness: How layers of innovation weave together to form multilayered tapestries of human cultures. *Mind & Society* 10: 79–96.
- Upal, M.A. 2011b. Memory, mystery, and coherence: Does the presence of 2–3 counterintuitive concepts predict cultural success of a narrative? *Journal of Cognition and Culture* 11: 23–48.
- Upal, M.A., L.O. Gonce, R.D. Tweney, and D.J. Slone. 2007. Contextualizing counterintuitiveness: How context affects comprehension and memorability of counterintuitive concepts. *Cognitive Science* 31: 415–439.
- Vidal, F., and B. Kleeberg. 2007. Introduction: Knowledge, belief, and the impulse to natural theology. *Science in Context* 20: 381–400.

- Vosniadou, S., and W.F. Brewer. 1992. Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology* 24: 535–585.
- Waytz, A., J. Capaccio, and N. Epley. 2010. Who sees human?: The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science* 5: 219–232.
- Welker, M. 2006. Science and theology: Their relation at the beginning of the third millennium. In *The Oxford Handbook of Religion and Science*, eds. P. Clayton and Z. Simpson, 551–561. Oxford: Oxford University Press.
- Whitehouse, H. 1992. Memorable religions: Transmission, codification, and change in divergent Melanesian contexts. *Man* 27: 777–797.
- Whitehouse, H. 1995. *Inside the Cult: Religious Innovation and Transmission in Papua New Guinea*. Oxford: Clarendon.
- Whitehouse, H. 2001. Transmissive frequency, ritual, and exegesis. *Journal of Cognition and Culture* 2: 167–181.
- Whitehouse, H. 2004a. *Modes of Religiosity: A Cognitive Theory of Religious Transmission*. Walnut Creek, CA: Alta Mira.
- Whitehouse, H. 2004b. Theorizing religions past. In *Theorizing Religions Past*, eds. H. Whitehouse and L.H. Martin, 215–232. Walnut Creek, CA: Alta Mira.
- Whitehouse, H. 2008. Cognitive evolution and religion: Cognition and religious evolution. In *The Evolution of Religion: Studies, Theories, & Critiques*, eds. J. Bulbulia, R. Sosis, E. Harris, R. Genet, C. Genet and K. Wyman, 31–41. Santa Margarita, CA: Collins Foundation Press.