BEHAVIORAL AND BRAIN SCIENCES (2006) **29**, 453–498
Printed in the United States of America

The folk psychology of souls

Jesse M. Bering
Institute of Cognition and Culture, Queen's University Belfast, Belfast BT7 1NN, United Kingdom.
j.bering@qub.ac.uk **qub.ac.uk/icc**
http://www.qub.ac.uk/schools/InstituteofCognitionCulture/Staff/JesseMBering/

Abstract: The present article examines how people's belief in an afterlife, as well as closely related supernatural beliefs, may open an empirical backdoor to our understanding of the evolution of human social cognition. Recent findings and logic from the cognitive sciences contribute to a novel theory of existential psychology, one that is grounded in the tenets of Darwinian natural selection. Many of the predominant questions of existential psychology strike at the heart of cognitive science. They involve: *causal attribution* (why is mortal behavior represented as being causally related to one's afterlife? how are dead agents envisaged as communicating messages to the living?), *moral judgment* (why are certain social behaviors, i.e., transgressions, believed to have ultimate repercussions after death or to reap the punishment of disgruntled ancestors?), *theory of mind* (how can we know what it is "like" to be dead? what social-cognitive strategies do people use to reason about the minds of the dead?), *concept acquisition* (how does a common-sense dualism interact with a formalized socio-religious indoctrination in childhood? how are supernatural properties of the dead conceptualized by young minds?), and *teleological reasoning* (why do people so often see their lives as being designed for a purpose that must be accomplished before they perish? how do various life events affect people's interpretation of this purpose?), among others. The central thesis of the present article is that an organized cognitive "system" dedicated to forming illusory representations of (1) psychological immortality, (2) the intelligent design of the self, and (3) the symbolic meaning of natural events evolved in response to the unique selective pressures of the human social environment.

Keywords: causal reasoning; death concept; evolutionary theory; existential psychology; folk biology; intelligent design; intentionality; mental representation; teleological reasoning; theory of mind

> Life is a jest, and all things show it; I thought so once, and now I know it.
>
> — John Gay, *Epitaph*

1. Introduction

By stating that psychological states survive death, one is committing to a radical form of mind-body dualism. Yet this radicalism is especially common. In the United States alone, 95% of the population reportedly believes in life after death (Greeley & Hout 1999; Lester et al. 2002). The majority of people from other societies, as well, see death as a transitional event that unbuckles the ethereal self from its body. The *soul* is typically represented as the conscious personality of the decedent and the once animating force of the now inert physical form (Thalbourne 1996). Although there are many varieties of afterlife beliefs, each – at least implicitly – shares a dualistic view of the self as being initially contained in bodily mass and as exiting or taking temporary leave of[1] the body at some point after the body's expiration (Bloom 2004; Boyer 2001).

There is clear evidence showing that emotive factors can be powerful contributors to people's belief in life after death (e.g., Alvarado et al. 1995; Dechesne et al. 2003; Thalbourne 1996). In general, psychologists who study this area have tended to focus on individual differences, specifically the role of death anxiety, and have posited a variety of "comfort hypotheses" involving the human motivation to construct such supernatural beliefs. In contrast, less is known about the basic components underlying the strong cognitive bias to entertain belief in an immortal soul (Astuti, forthcoming a). These more basic questions concerning the cognitive architecture behind afterlife representations are also important pieces of the puzzle and will be explicitly addressed in the present article. Whatever one's personal motivations for rejecting or endorsing the idea of an immaterial soul that can defy physical death, the ability to form any opinion on the matter would be absent if not for our species' defining capacity to differentiate unobservable minds from observable bodies (Povinelli & Bering 2002; Suddendorf & Whiten 2001; Tomasello & Call 1997).

Some researchers have already begun laboratory investigations into the question of whether humans are "common sense dualists," work that seems to have implications for our understanding of people's intuitive conceptions of souls and the afterlife (see Bloom 2004). For example, in a modification of the classic expectancy violation paradigm (which uses looking time as a measure of nonverbal infants' "surprise" at an event), Kuhlmeier et al. (2004) positioned identical twin experimenters at different points in the laboratory to test 5-month-olds'

BERING JESSE, Reader in the School History and Anthropology Departments and Director of the Institute of Cognition and Culture, Queen's University Belfast, has focused his work on the psychological foundations of supernatural belief. An experimental cognitive scientist, he is also a specialist in the evolutionary bases of religion.

 0140-525X/06 $12.50

ability to reason about the law of continuous motion as it applies to human bodies. Like any material substance, human bodies cannot go from A → C without first passing along the trajectory of B (a contiguous space between the two points). For inanimate objects, infants are surprised (i.e., look longer) when the object disappears from behind one barrier and then seems to reemerge from behind another nonadjacent barrier. In the case of a human who violates the law of continuous motion, however, 5-month-olds are not surprised (i.e., they do not look longer at this event than the non-violation event). The authors speculate that "infants do not readily view humans as material objects" (Kuhlmeier et al. 2004, p. 101) and that an "appreciation that people *are* just objects may be a developmental accomplishment" (p. 102; emphasis in original).

But how do we get from the common-sense dualism of infants to beliefs of the afterlife so soberly entertained by adults? Even a superficial pass over such beliefs strikes one as involving many of the core problems in cognitive science: *causal attribution* (how is mortal behavior causally related to one's afterlife? how are dead agents envisaged as communicating messages to the living?), *moral judgment* (why are certain social behaviors, i.e., transgressions, believed to have ultimate repercussions after death or to reap the punishment of disgruntled ancestors?), *theory of mind* (how can we know what it is "like" to be dead? what social-cognitive strategies do people use to reason about the minds of the dead?), *concept acquisition* (how does a common-sense dualism interact with a formalized socio-religious indoctrination in childhood? how are supernatural properties of the dead conceptualized by young minds?), *teleological reasoning* (why do people so often see their lives as being designed for a purpose that must be accomplished before they perish? how do various life events affect people's interpretation of this purpose?), and so on.

In what follows, I examine how this human folk psychology of souls, as well as closely related supernatural beliefs, may open an empirical backdoor to our understanding of the evolution of human social cognition. Recent findings and logic from the cognitive sciences contribute to a novel theory of existential psychology, one that is grounded in the tenets of Darwinian natural selection. The central thesis of the present article is that an organized cognitive "system" dedicated to forming illusory representations of (1) psychological immortality, (2) the intelligent design of the self, and (3) the symbolic meaning of natural events evolved in response to the unique selective pressures of the human social environment.

2. Psychological immortality as a cognitive default

> I'm a materialist, I swear it to you; I'm not going crazy. But something's the matter. I see my corpse; that's not hard but I'm the one who sees it, with my eyes. I've got to think . . . think that I won't see anything anymore and the world will go on for the others. We aren't made to think that.
>
> — Jean-Paul Sartre (1937/1969), *The Wall: And Other Stories*

From an evolutionary perspective, it is important to first ask whether humans "naturally" reason about death as a transitional state of consciousness or simply acquire such ideas through cultural exposure (perhaps from adults who "invent" such notions to ameliorate their own death anxiety; see, e.g., Dechesne et al. 2003; Harris & Giménez 2005). Although conventional wisdom tends to favor a general learning hypothesis for the origins of afterlife beliefs, recent findings suggest a more complicated developmental picture.

For example, in a study by Bering and Bjorklund (2004), children (as well as an adult comparison group) were presented with a puppet show in which an anthropomorphized mouse was killed and eaten by an alligator, and then asked about the biological and psychological functioning of the now-dead mouse. Kindergartners understood that various biological imperatives (e.g., the capacity to be sick, the need to eat, drink, and relieve oneself) no longer applied to the dead mouse. The majority of these children even said that the brain of the dead mouse no longer worked, which is especially telling given that children at this age also understand that the brain is "for thinking" (Bloom 2004; Gottfried & Jow 2003; Johnson & Wellman 1982; Slaughter & Lyons 2003). Yet when asked whether the dead mouse was hungry or thirsty, or whether it was thinking or had knowledge, most kindergartners said yes. In other words, young children were cognizant of the fact that the body stops working at death but they viewed the mind as still active. Furthermore, both the children and adults were particularly likely to attribute to the dead mouse the capacity for certain psychological states (i.e., emotions, desires, and epistemic states) over others (i.e., psychobiological and perceptual states), a significant trend that will be addressed in the following section.

In general, however, kindergartners were more apt to make psychological attributions to the dead mouse than were older children, who were not different from adults in this regard. This is precisely the opposite pattern that one would expect to find if the origins of such beliefs could be traced exclusively to cultural indoctrination. In fact, religious or eschatological-type answers (e.g., Heaven, God, spirits, etc.) among the youngest children were extraordinarily rare. Thus, a general belief in the continuity of mental states in dead agents seems not something that children acquire as a product of their social–religious upbringing, because increasing exposure to cultural norms would increase rather than attenuate afterlife beliefs in young minds. Instead, a natural disposition toward afterlife beliefs is more likely the default cognitive stance and interacts with various learning channels (for an alternative interpretation, see Astuti, forthcoming a). Moreover, in a follow-up study that included Catholic schoolchildren, this incongruous pattern of biological and psychological attributions to the dead mouse appeared even after controlling for differences in religious education (Bering et al. 2005).

Unlike intuitive reasoning about dead agents' bodies, which may help to motivate physical avoidance of these dangerous objects in the environment (via the emotion of disgust or agency detection mechanisms which err on the side of caution for ambiguously dead/sleeping agents; Barrett & Behne 2005; Rozin et al. 1993), intuitive reasoning about dead agents' minds would seem to leave open the possibility for continued social relationships with the dead.

2.1. The simulation constraint hypothesis and the afterlife

> Our own death is indeed unimaginable and whenever we make the attempt to imagine it we can perceive that we really survive as spectators.
>
> — Sigmund Freud, *Thoughts for the Times on War and Death*
>
> Try to fill your consciousness with the representation of no-consciousness, and you will see the impossibility of it. The effort to comprehend it causes the most tormenting dizziness. We cannot conceive of ourselves as not existing.
>
> — Miguel de Unamuno (1912/1954), *Tragic Sense of Life*

The causal mechanisms that lead young children to represent dead agents' minds as being psychologically active have yet to be precisely identified. Nevertheless, there is evidence that simulation constraints (i.e., the inability to know what it is "like" to be dead) may comprise an important set of factors. Like reasoning about one's past mental states during dreamless sleep or while in other somnambulistic states, consciously representing a final state of unconsciousness poses formidable, if not impassable, cognitive constraints (Barrett 2004; Bering 2002a; Bering & Bjorklund 2004; Bering et al. 2005; Clark 1994; Gilbert 2001; Nichols, in press). By relying on simulation strategies to derive information about the minds of dead agents, one would be compelled to put themselves "into the shoes" of such organisms, which is an impossible feat. These constraints may lead to a number of telltale errors, namely "Type I" errors (inferring mental states when in fact there are none), regarding the psychological status of dead agents. Koocher (1973, p. 374) described, for instance, how a group of children tested on death comprehension reflected on what it might be like to be dead "with references to sleeping, feeling 'peaceful,' or simply 'being very dizzy'."

Attempts to simulate dead agents' minds may even result in Type I errors made by adults who profess not to believe in the afterlife. Bering (2002a) found that when undergraduate students were asked to reason about the psychological abilities of a protagonist who had just abruptly died in an automobile accident, even some participants who later classified themselves as "extinctivists" (i.e., those who endorsed the statement "what we think of as the 'soul,' or conscious personality of a person, ceases permanently when the body dies"; after Thalbourne 1996), nevertheless stated that the dead person *knew* that he was dead.

In addition, there is reason to believe that certain types of mental states are more difficult to imagine being permanently without than are others. In the study by Bering and Bjorklund (2004), for example, participants at every age were more likely to attribute emotions, knowledge, and desires to the dead mouse than that they were psychobiological and perceptual states (see also Bering et al. 2005). This may be understood in relation to children's growing scientific knowledge. With regard to psychobiological states, such as hunger or thirst, Slaughter and her colleagues have shown that once children display an understanding of the vitalistic purpose of the behaviors tied to these states (i.e., that eating and drinking function to sustain life), this knowledge facilitates scientific reasoning about death (Slaughter & Lyons 2003; Slaughter et al. 1999). Indeed, children who appeal to a vitalistic biological framework when reasoning about human bodies are more precocious in their understanding of death (Slaughter & Lyons 2003).

Similarly, because perceptual states are closely tied to obvious bodily structures, children who possess teleofunctional biological knowledge about these structures (e.g., that ears are "for hearing") may begin to reason that, so long as the body has stopped functioning, the capacity for such states must also become defunct at death (O'Neill & Chong 2001).

In addition, because individuals are aware from their own previous or current experiences what it is like, say, *not* to be sleepy, *not* to hear, or *not* to be hungry, they may draw from the phenomenal negation of such states and apply these experiences to the minds of dead agents. Thus, in some cases, simulation may actually corroborate scientific knowledge and further reduce Type I errors.

In contrast to these categories of psychological states, however, the nature of the body's role in producing the subjective experiences of emotions, desires, and beliefs seems not as amenable to children's scientific theories of dead minds (or, indeed, even to adults' formulation of scientific theories regarding phenomenal consciousness and the brain, e.g., *qualia*; see McGinn 1991). These aspects of consciousness are not obviously related to the body's survival, nor are they linked to external bodily accoutrements (i.e., sense organs) that become "broken" by death. In the absence of scientific theory concerning the isomorphic relationship between the brain and the mind, individuals may defer to a simulation strategy in reasoning about dead others, a strategy that inevitably leads to Type I errors for these particular mental capacities (Bering 2002a; Clark 1994; Gilbert 2001; Nichols, in press). Firsthand experiences with the phenomenal negation of mental functions such as desires, emotions, and thought can never be had because these states are constantly "turned on" during conscious periods (e.g., it is epistemologically impossible to know what it is like not to think), making people inclined to impute these capacities to dead agents. Indeed, in looking at participants' response latencies to state that a dead protagonist lacked the capacity for various mental states, Bering (2002a) reported that it took people longer to answer that this was the case for "difficult-to-imagine-the-absence-of" states (e.g., desire: "Now that he's dead, does he *want* to be alive?") than for "easy-to-imagine-the-absence-of" states (e.g., psychobiological: "Now that he's dead, is he still *sleepy*?").[2]

2.2. Offline social reasoning: Why the afterlife is a place

> I forced myself to stop thinking of her as someone still somewhere, if only in memory, still obscurely alive, breathing, doing, moving, but as a shovelful of ashes; as a broken link, a biological dead end, an eternal withdrawal from reality.
>
> — John Fowles (1978), *The Magus*

In addition to simulation constraints, there are other aspects of human social cognition that may encourage attributions of continued psychological functioning to dead agents. When investigating peoples' intuitive conceptions of dead agents' minds, we are wise to remember, for instance, that human relationships are largely characterized by offline social events; those with whom we have relationships are only periodically directly

observable (e.g., Dunbar 1993; 2004). An offline social system leads us to tacitly assume that individuals with whom we have relationships are engaged in actions even when we cannot observe them doing so. The fact that your mother is not in the room at the moment does not compromise your capacity to reason about her mind, though obviously the accuracy of your social judgments will be limited. When conjuring up her offline image you are likely to imagine her as *somewhere* and as doing *something* – in the kitchen washing dishes, in bed sleeping, playing squash with the neighbor, and so on. Similarly, the dead are envisaged not as inanimate objects slowly decomposing in situ under the earth, but instead as having relocated to some unobservable locale where they are very much "living" their dead lives.

When it comes to death, human cognition apparently is not well equipped to update the list of players in our complex social rosters by accommodating the recent nonexistence of any one of them. This is especially the case, of course, for individuals who have played primary roles in our social lives, who did so for a long time, and who were never presumed to be continuously stationary when they were out of our sight. Because our minds are designed for offline as well as online social processing, we expect the periodic physical absence of social partners. Casual observation reveals that individuals will often, for example, pick up the phone with the intention of calling the decedent or fleetingly imagine how the decedent will react when told about some good news, only to remember that the person is not where they usually are – they have "passed on" to someplace else.

Although these automatic cognitions are probably the residue of habitual social behaviors, they also reveal something about the challenges faced by the human cognitive system when it attempts to process information concerning the truth about dead agents' *physical* whereabouts. A person who has recently died and whose body has already been disposed of may continue to be processed by an offline social system for an undetermined period of time. This place error is seemingly compounded by nonnegotiable simulation constraints that tempt us into reasoning about these dead agents' continued psychological functioning (as discussed in the previous section).

2.3. By-product versus functional analyses of belief in immortal souls

> THE CHILD: I'm frightened.
> THE WOMAN: And so you should be, darling. Terribly frightened. That's how one grows up into a decent, god-fearing man.
>
> — Jean-Paul Sartre (1937/1969), *The Flies*

There may be good reason to argue that natural selection operated on the foregoing psychological biases. Representations of the afterlife are culturally recurrent, proximally driven by emotions, frequently implicated in social and reproductive matters, and superficially fitted to the ecological niche in which the human organism develops (Bering & Bjorklund 2004; Dechesne et al. 2003; Reynolds & Tanner 1995). These features are consistent with what we know about the nature of psychological adaptations (Bjorklund & Pellegrini 2002; Tooby & Cosmides 1992). This is not to say that *specific* afterlife beliefs – the culturally variable vicissitudes of the hereafter – are direct products of natural selection. As will soon be discussed, investigators such as Boyer (2001) and Atran (2002) have shown that the "selection" of explicit religious ideas occurs at the cultural level, with the "survival" of such ideas being a feature of their ability to become ensconced in the evolved architecture of exposed human minds (Sperber & Hirschfeld 2004). Instead, it is to argue that the subtle contours of a uniquely human adapted design may stand out when closely examining the folk psychology of souls, an intuitive pattern of reasoning that does not appear to hinge on the presence of explicit religious concepts per se (Bering 2002a).

But how might a representational bias for envisioning personal immortality have impacted the net genetic fitness of individual humans in ancestral environments? Unfortunately, among cognitive scientists, scant attention has been paid to the evolutionary significance of the human capacity to represent the self *sub specie aeternitatis* ("under the aspect of eternity"). Instead, many scholars categorize afterlife concepts in the same way they do other types of religious concepts, as especially virulent strains of culturally transmitted ideas that are highly effective at pirating core cognitive architecture (Atran 2002; Boyer 2001; for an exception, see Baron-Cohen 1999). According to this perspective, only the cognitive architecture itself can be the product of natural selection; religious ideas are seen as simply being parasitic on this evolved architecture – as nothing more than noise that shares a general frequency between cultures (e.g., Pyysiäinen 2001; Sperber & Hirschfeld 2004). For example, in his book *Religion Explained*, Boyer (2001, p. 40) writes: "People have religious notions and beliefs because they acquired them from other people. On the whole, people get their religion from other members of their social group."

Boyer and other *cultural epidemiologists'* view afterlife concepts, as well as other types of supernatural concepts, as unavoidable carryovers of cultural selection. Specifically, Boyer (2000; 2001) has argued that religious ideas exploit information-processing mechanisms into paying attention to them because they violate ontological regularities by hybridizing or transgressing natural categories (see also Mithen 1996). Thus, religious ideas are especially likely to attach to evolved cognitive templates that are designed for reasoning about exemplars from natural categories – such as PERSON or ANIMAL – because these templates act as flypaper for salient, "counterintuitive" cases (Atran & Norenzayan 2004; Barrett 2000; 2004; Pyysiäinen 2001; Slone 2004; Sperber & Hirschfeld 2004). According to Boyer (2003a), then, a ghost is a person who is without a physical body and as such is a conceptually seductive idea. The concept of an afterlife therefore is easily generated and transmitted between minds. Like all other religious concepts, however, it is otherwise treated as a biologically sterile by-product.

Similarly, Sperber and Hirschfeld (2004, p. 44) write that, "explaining religion by a religious disposition lacks insight and plausibility." Instead, these scholars argue that religion is a non-adaptive by-product (i.e., a spandrel), one that arises through adapted human cognition acting in concert with culturally migrating counterintuitive concepts that change shape both within and between minds (for critical reviews, see Alcorta & Sosis 2005; Bulbulia 2004; 2005; Sosis & Alcorta 2003).

2.4. Inhibition and the preservation of reputation

I believe that I am in hell, therefore I am.

—Arthur Rimbaud (1873/1999), *A Season in Hell*

Once the ability to entertain supernatural agent concepts evolved, such ideas might have led our ancestors to inhibit socially proscribed actions out of the fear that gods or dead agents, now "full access strategic agents," were watching them (Boyer 2001). Some empirical support for this general argument was found in a recent study by Bering et al. (2005). In this study, undergraduate students who were casually told that a ghost was recently spotted in the laboratory were less willing to cheat on a competitive computer task – as measured by latency of response to delete the "accidentally" revealed answer – when they were tested alone in the room than were control participants who heard nothing of the fictitious ghost (see also Burnham & Hare, in press; Haley & Fessler 2005). In the case of supernatural beliefs, therefore, it is helpful to highlight Haselton and Buss's (2003, pp. 29–30) general adaptationist point that, "the human mind is designed to reason adaptively, not truthfully or even necessarily rationally." The genetic fitness effects of such behavioral inhibition have real currency in natural selection theory.

The relationship between supernatural morality and behavioral inhibition is potentially a very important point for evolutionary biologists. Because natural selection is pragmatic, the illusion of a supernatural morality, if it served to curb selfish behaviors and thus preserved social reputation in the ancestral past, may be an illusion by design (Alcorta & Sosis 2005; Bering 2005; Bering & Johnson 2005; Bering et al. 2005; Boyer 2001; Bulbulia 2004; Dunbar 2004; Hinde 1999; Johnson & Krüger 2004). Many writers have argued that, at some point in the recent evolutionary past, hominid sociality underwent a relatively abrupt shift that was characterized by strong selective forces operating on reputation-related behaviors (Alexander 1987, p. 110; Bering & Bjorklund, in press; Bering & Shackelford 2004; Daly & Wilson 1994; Emler 1994; Frank 1988; Goffman 1959, 1963; Hilton et al. 1993; Schelling 1960; Wright 1994). Because of the risks associated with social detection of selfish acts, and the peculiar "stickiness" of bad reputations (e.g., Baumeister et al. 2001; Goffman 1963), psychological traits that facilitated the inhibition of selfish acts were likely subjected to natural selection. The costs of underestimating the risk of social detection would have been disproportionately greater than the costs of prosocial decisions that were contextually maladaptive. Even if altruism was costly *every* time, if it avoided a lethal cost *once*, those other costs would become negligible (Nettle 2004).

Experimental findings of prosocial behavioral change in light of supernatural primes (e.g., Bering et al. 2005) also link up with the ethnographic database concerning afterlife beliefs. In some religious ideologies, the fate of the soul *after death* is determined by the social behaviors of the individual *during life*. Reflections on the ultimate consequences of (im)moral actions (e.g., whether the soul is expelled to Hell or dissipated in nirvana) should be capable of exerting a causal influence on today's overt behavior, which would have the effect of preserving reputation by encouraging the inhibition of selfish acts or facilitating self-control. In the United States, for example, the majority (79%) of people believe that there will come a day when God judges them and decides whether they will go to Heaven or Hell (Gallup Organization 1999; see also Lester et al. 2002). For current purposes, such poll data may actually be misleadingly low. What people say they believe about the supernatural and how they implicitly reason are quite different things (Subbotsky 1997; 2001). Scientific knowledge about causal relations between behaviors and consequences may therefore destabilize this adapted system, but more by overriding supernatural beliefs than by replacing them (McCauley 2000; Subbotsky 2001).

In many traditionalist religious societies, the emphasis is on worldly punishment for moral transgressions, whereby norm violators are visited by sickness, poverty, or other types of misfortune (see Bering & Johnson 2005). Mostly, punishment is seen as being imposed by disgruntled ancestors (Hinde 1999; Reynolds & Tanner 1995). In some cases, belief in the vicarious punishment of dead agents achieves similar prosocial effects. In medieval Europe, where people's social behaviors were thought to determine the fate of dead loved ones whose souls were at limbo in purgatory, thoughts of the dead were so prevalent in the daily affairs of the living that at least one historian has even referred to the dead as constituting a separate "age group" (Davis 1977; as cited in Luria 2001).

Although critical developmental studies have yet to specifically address the etiology of full-access strategic agent concepts and their consequences for behavioral inhibition, some related findings with young children do point to a human cognitive system prepared to reason about "omniscient" supernatural agents (Bering 2005; Bering & Johnson 2005). In a recent study, Barrett et al. (2001) report that because, theologically, God is all-knowing and therefore cannot hold false beliefs (and therefore cannot be deceived), the social cognitive systems of young children may be better suited to reasoning about the culturally postulated mind of God than about the epistemologically limited minds of humans and other animals. For example, whilst 3-year-olds *incorrectly* reason that a naïve person knows the true contents of an inaccurately labeled box, they *correctly* reason (at least, in a theological sense) that God knows the true contents as well. Thus, according to the authors, because of egocentric biases in early childhood, there may be cognitive precursors for full access strategic agent concepts that developmentally precede even natural mental agent concepts.

3. Souls and intelligent design

The concept of man in the mind of God is comparable to the concept of paper-cutter in the mind of the manufacturer, and, following certain techniques and a conception, God produces man, just as the artisan, following a definition and a technique, makes paper-cutter. Thus, the individual man is the realization of a certain concept in the divine intelligence.

— Jean-Paul Sartre, *Existentialism and Human Emotions*

However, to understand the relationship between belief in gods or other supernatural agents who are interested in our social behaviors and belief in immortal souls requires that we disentangle several related strands of causal reasoning. Consider that if God does not exist, then the

unique self (i.e., the individual "soul" of any given person) cannot be the product of intelligent design; rather, it is simply the end product of standard machinations of genetic and environmental recombination. If the soul is not the product of intelligent design, then there is no teleological function that it is designed to fulfill, no *raison d'être* to explain its existence beyond human attributions of purpose. The task remains for cognitive scientists to determine why the teleological position is so frequently adopted, and prospers so vehemently, over the mechanistic alternative. The human mind cannot seem to easily accommodate itself to a godless, evolutionary canon when it comes to the self's existence.

In fact, resistance to the mechanistic theory of natural selection may have as much to do with a cognitive bias toward intentionality as it does with an emotionally laden or moralistic bias. Recent findings converge to show that humans have a strong teleological bent when it comes to reasoning about the origins of artifacts, animals, and natural objects (e.g., Evans 2001; German & Barrett 2005; Kelemen 2004; Kelemen & DiYanni 2005). Kelemen (2004) has even gone so far as to dub children "intuitive theists" because of their so-called "promiscuous teleology." According to Kelemen, most young children would prefer the teleo-functional explanation that a cloud is "for raining" rather than assent to the experimenter's suggestion that perhaps raining is just something that a cloud does. This cognitive bias shows that young children are "endorsing the view that natural entities are 'made for something' and that is why they are here" (Kelemen 2004, p. 295). In a similar vein, Evans and her colleagues have found evidence that most young children prefer creationist arguments over evolutionary ones when reasoning about the origins of species (e.g., Evans 2001).

Teleological reasoning is often applied to the origins of the self, as well, such as talk about what one was "born to do" or that one is leading a life that he or she was not "meant for." Indeed, the term *conceive* (from the Latin *concipere*, "to take in and hold"), though originally used to describe impregnation ("to take into the womb, become pregnant"), was within that same century (c. 1280–1340) adopted to describe an intentional mental process ("to take into the mind").

The tendency for people to reason about the special purpose of the unique self may differ from other forms of teleo-functional reasoning in two important ways. First, it appears to be much more resistant to scientific knowledge. Although teleo-functional beliefs about natural objects are found in Romanian Gypsy adults, a group that does not possess scientific knowledge regarding natural artifact origins (see Kelemen 2004), they generally decline with age and are relatively rare among scientifically educated adults. In contrast, ascriptions of intelligent design when reasoning about the purpose of individual lives appear to remain stable (and perhaps even increase) over the life course, probably due to the accrual and retrospective interpretation of autobiographical experience (Bering 2003b; Bruner 2001; McAdams 2001).

Second, when it comes to lay beliefs about souls, attributions of purpose occur frequently for individual members of the same conceptual family. People tend to ascribe special purpose more often to the specific case – such as "what am *I* meant for?" – than they ascribe shared purpose to members of the same natural category – such as "what is the *human species* meant for?" For no other natural categories do such special teleological ascriptions seem to occur. (Imagine an evolutionary biologist hypothesizing about the special purpose of a specific heart of a specific organism of all the possible such organisms within a given species.) Even for artifacts, teleo-functional judgments for class categories (e.g., CHAIR) appear to trump within-category exemplars (e.g., both a gothic revival style and a Chinoisserie style chair are "for sitting" although they may differ in design for posturing the body), and rarely occur within the exemplar class itself (e.g., the special purpose of an individual Chinoisserie style chair) (Defeyter & German 2003).

The categorical question "Why am I here?" is important for evolutionary analysis because it may set the stage for an obligatory social relationship between the self and its presumed supernatural creator. If this cognitive illusion, enriched with social affect, plays a causal role in generating genetic fitness-enhancing responses (e.g., through the individual's behavioral compliance with moral norms which the creator is believed to have authored), then an adaptationist hypothesis for the folk psychology of souls gains support.

The tendency to endow human lives with an a priori meaning is particularly obvious in the wake of recent loss. Despite differences in religiosity, individuals who are in mourning commonly report feeling a sense of meaninglessness (Davis & Nolen-Hoeksema 2001; Golsworthy & Coyle 1999; Smith et al. 1992; Yalom 1980). Such existential despair, characteristic of the early stages of the grieving process, betrays people's implicit belief that they are part of a privileged social relationship with some abstract agent who exerts a causal influence over their everyday lives. Many types of "premature" death (e.g., accidents, fatal illnesses, homicides) seem to force surviving individuals to acknowledge that this privileged social relationship is illusory: the existence of the self is abruptly surrendered to a veridical belief in the fundamental and mindless laws of natural probability. The resulting existential despair can be attributed to the realization that the predictability and controllability of one's own death, like that of the decedent's, is in actuality very low.

In this light, there is no emotionally invested God who favors or disfavors the continued survival of the self. Consequently, whatever social contracts previously entered into with this nonexistent agent that led the self to expect a *reasonable* deferment of death until old age are exposed as being spurious. Avenues by which individuals may reenter into this illusory contract include "just world" beliefs (e.g., by reasoning that the person must have been somehow deserving of death), and judging that the decedent was different from themselves (and thus unlikely to have been in the same sort of privileged social relationship with God) (e.g., Hafer & Bègue 2005; Lerner 1980; Lerner & Miller 1978; Pyszczynski et al. 1995).

The fact that most individuals do tend to reenter into these illusory social contracts, even in the face of seemingly egregious violations, suggests that the affective push to do so is capable of overthrowing any rationalist *Weltanschauung*. However, if this is the expression of an evolved system, as the current model alleges, then any explicit philosophical position that discards meaning is naturally disadvantaged, because the self can no sooner

"choose" to be a subjective atheist than retinas can "choose" not to convert light energy into signals that are carried to the brain by the optic nerve (Bering 2005; McCauley 2000). Adapted psychological systems, by definition, determine the way that information can be processed due to design solutions in the brain that were engineered by natural selection. In the present case, just as we can close our eyes to prevent light from being converted into neural signals, science may provide a minority (i.e., nonbelievers) with the armamentarium to close their eyes to the supernatural.

3.1. Suicide as a violation of intelligent design

> I condemn that nature which, with such impudent nerve, brought me into being in order to suffer – I condemn it in order to be annihilated with me.
>
> —Fyodor Dostoyevsky (1877/1949), *Diary of a Writer*

Public opinions concerning suicide further serve to highlight the role of the design stance in the existential domain. Those who believe that one's life is owned by God are more likely to view suicide – as well as abortion, capital punishment, and medical euthanasia –as being morally wrong (Ross & Kaplan 1994; Worthen & Yeatts 2001). It is not suicide per se that sheds light on this teleo-functional bias, but the moral repugnance for the act. Religious rules against suicide reveal a more complex cognitive stance than is immediately apparent. It is a counterintuitive stance in which the self's will is seen as imposing itself over the will of the creator of the self's will. According to the premise that a person's life belongs to God, an individual does not have the right to purposefully cause his or her own death, because this right is seen as being God's alone. This conception suggests that suicide is viewed as a moral transgression in which an individual "cheats" God by stealing the latter's power of intentionality in causing the self's death. Suicide therefore becomes a form of intellectual theft; the self redesigns its end in an act of mutiny against its creator.

Suicide must be distinguished from acts of martyrdom, in which an individual engages in self-sacrifice as a political or wartime strategy (e.g., "suicide bombers" or kamikaze pilots) (Atran 2003). Even here, however, we see how intentionality critically underlies the folk psychology of souls. Although suicide is treated as a sin by many of the world's religions, including Islam, those who are martyrs are seen by some religious adherents as having been chosen by God to fulfill His wishes and as being rewarded with special experiential luxuries in the afterlife. For example, during World War II, one of the most intense and successful military operations ever launched by Japanese fighter pilots against an American fleet was deemed "Operation Heaven" by the Japanese commander (Blanchard n.d., p. 17), and "kamikaze" is literally translated as "divine wind."

The religious messages that are conveyed by charismatic leaders may be especially seductive because they capitalize on an innate teleological bias for ascribing a special purpose to the unique self. For example, in a transcribed television interview from CNN in 1997, Osama Bin Laden commented that: "We believe that no one can take out one breath of our written life as ordained by Allah. We see that getting killed in the cause of Allah is a great honor wished for by our Prophet." When juxtaposed with simulation constraints concerning what the afterlife may be "like" for those who sacrifice themselves for prosocial in-group reasons, this becomes a particularly volatile social cognitive phenomenon since martyrs are promised privileged states of consciousness after death. As one member of the Palestinian group Hamas put it: "By pressing the detonator, you can immediately open the door to Paradise – it is the shortest path to Heaven" (Hassan 2001).

4. Meaning, morality and the afterlife

> Some say that we shall never know and that to the gods we are like flies that the boys kill on a summer day, and some say, on the contrary, that the very sparrows do not lose a feather that has not been brushed away by the finger of God.
>
> —Thornton Wilder, *The Bridge of San Luis Rey*

Because the adjudication of an afterlife of eternal reward or damnation is seen as superseding "mere" human authority, people's understanding of the *origins* of moral deontology – what one *ought* and *ought not* to do in life – shows a strong cognitive bias toward belief in a supernatural creator of human morality (rather than, for instance, a bias toward belief in design by nature or human whim). Reincarnation beliefs that rely on karmic principles are no exception because such principles require an intelligent designer of this morality-based rebirthing cycle. This overall vein of reasoning helps to explain why people expect divine retribution for moral transgressions only, rather than, say, for breaches of social etiquette (e.g., Roes & Raymond 2003). As Camus (1943/1991) writes, "revolt against men is also directed against God" (p. 94). From a genetic fitness perspective, what is important is that it is moral transgression that scars reputation most deeply and has the most costly effect on future social relations (Goffman 1963) and therefore behavioral compliance in this domain is critically important.

Surprisingly, cognitive scientists who study religion have given the topic of morality relatively short shrift. For example, Atran and Norenzayan (2004) recently argued that culturally acquired supernatural concepts (cf. Boyer 2001) receive emotional staying power because they are lent support by an evolved *hyperactive agency detection device* (see also Atran 2002; Barrett 2000; Guthrie 1993). According to Atran and Norenzayan, this mechanism serves the protective function of hyper-vigilance in potentially dangerous environments, but as a consequence, affectively primes individuals and causes them to overattribute intentions to the natural world, such as might happen when a branch falls in the forest. The authors conclude that "supernatural agents are readily conjured up because natural selection has trip-wired cognitive schema for agency detection in the face of uncertainty" (Atran & Norenzayan 2004, p. 720). In particular, supernatural attributions occur because environmental stimuli "achieve the minimal threshold for triggering hyperactive facial-recognition and body-movement recognition schemata that humans possess" (p. 720).

Atran draws on findings from developmental psychology showing that agency overgeneralization is an innate

feature of human cognition. For instance, in a variety of controlled experiments using nonverbal measures, Csibra and his colleagues have demonstrated that, if causal cues indicating rational agency are present, even infants see inanimate movement as purposive behavior (e.g., 12-month-olds ascribed intentions to dots on a computer screen moving about in a "rational" manner; see Gergely & Csibra 2003).

Despite minor theoretical differences with Boyer's evolutionary model of religion, Atran (2002) is united with Boyer and other cultural epidemiologists in denying that religion is an adaptation. However, although the explanatory utility of cultural epidemiology theory has been unrivaled among recent attempts to explain the evolutionary basis of religion, it has problems of its own. By focusing on the role of concept acquisition, this work may be overshadowing more fundamental questions about the natural foundations of religion – questions raised in section 3 (Souls and intelligent design). For example, Atran and Norenzayan's (2004) model fails to account for people's tendency to assume that supernatural agents are responsible for traumatic life events (Deridder et al. 1999; McAdams 2001; Pepitone & Saffioti 1997; Weeks & Lupfer 2000). How can reasoning about the supernatural causes of, say, suffering a miscarriage, being felled by disease, or losing a loved one in an accident be triggered by facial-recognition and body-movement recognition schemata? There are no such environmental cues capable of breaking the "hair trigger" of the authors' proposed sensory driven hyperactive agency detector, yet supernatural attribution occurs (arguably even more so than for the exemplar hair-trigger cases).

Another approach to solving the riddle of religion is to address whether the self's view that it is something more than a material body subject to the mindless and amoral laws of nature is a product of natural selection (Dennett 1991; 1995). Evolutionary scholars in this area might then begin to shift the primary theoretical frame from one that centers on concept acquisition and agency detection to one that centers on models of self representation, morality, and meaning (Bering 2002b; 2003b). Although these approaches likely reflect complementary levels of analysis rather than alternative theoretical models, the cultural epidemiology approach has, to date, not successfully bridged the representation of supernatural concepts with the Darwinian currency of behavior. What is required to bridge this gap is the self, a conspicuously absent entity in the cognitive science of religion.

4.1. "Signs": Ascribing meaning to natural events

> The intentional stance is the strategy of interpreting the behavior of an entity (person, animal, artifact, whatever) by treating it as if it were a rational agent who governed its "choice" of "action" by a "consideration" of its "beliefs" and "desires" . . . the basic strategy of the intentional stance is to treat the entity in question as an agent, in order to predict –and thereby explain, in one sense – its actions or moves.
>
> —Daniel Dennett (1996), *Kinds of Minds*

If people naturally endow the events of their lives with a hidden purpose, the self may then hold expectations about the "behaviors" of supernatural causal agents, canonical expectations that conform to standard rules of fairness and justice. Research on just-world beliefs shows that people indeed operate under the assumption that others "get what they deserve," especially when they have little control over negative outcomes and when help cannot be meted out to unfortunate innocents (for a recent review of this literature, see Hafer & Bègue 2005). Although just-world researchers have not generally sought to interpret related religious notions, often implicit in this type of causal reasoning about fortune and misfortune is the idea that some behavior in the moral domain is connected to an unrelated, uncontrollable life event. Therefore, a central question is "who" is represented as tallying up our deeds and as meting out just deserts in the form of positive and negative life events (in whatever ontological domain these happen to be administered).

Bruner (1990) has argued that, in everyday social psychology, individuals will search for meaning whenever others' behaviors violate their expectations, or otherwise fail to adhere to sociocultural scripts. For example, subtle breaches of conversational maxims, such as non sequiturs or other types of "conversational implicatures," will often encourage a search of the partner's intentions (see also Baldwin & Moses 1996; Baron-Cohen et al. 1997). Whenever unexpected autobiographical events occur, individuals may similarly seek to identify the intentions of the supernatural agent who has caused these events (or at least allowed them to happen), because this is presumably a *purposeful* agent who adheres to unwritten rules of social reciprocity (Bering 2003b). In Nazi Germany, for instance, some Holocaust survivors reportedly thought that God had gone insane, since clearly he had breached the most basic of social agreements with his followers (see Wiesel 1961).[3]

This belief in a just world is so strong, in fact, that among many groups personal calamities and hardships are taken as evidence that the individual must have done something horribly wrong. Often the only suitable remedy for these hardships is spiritual excision by way of public confession. Among the Igbo of Nigeria, for example:

> [a]dultery by a wife is regarded as bringing supernatural punishment upon herself and her husband . . . thus if a woman experiences difficult labor, it is assumed that she has committed adultery and she is asked to give the name of her lover in order that the child be born. If a man falls sick, his wife may be questioned as to whether she has committed adultery. (Ottenberg 1958/1980, p. 124)

There may also be "nonreligious" developmental precursors to this moralistic interpretation of uncontrollable negative events. Piaget (1932/1965) argued that young children evidence a belief in *immanent justice* in which "the child must affirm the existence of automatic punishments which emanate from things themselves" (p. 251). Thus, in a classic study, Piaget (1932/1965) presented children aged 6–12 years with the story of a child who steals or disobeys and then, upon crossing a bridge, falls into the water when the bridge collapses. Nearly all (86%) of the youngest children in the study reasoned that the accident would never have happened were it not for the character's earlier misdeeds.[4]

Indeed, people who have violated some moral rule often appear expectant of existential punishment, and those guilty parties who find themselves untarnished by their wrongdoing may feel as though their current happiness is undeserved. This is another common theme in

literature, exemplified by the works of Victor Hugo (e.g., *Les Miserables*) and Fyodor Dostoyevsky (e.g., *The Brothers Karamazov, Crime and Punishment*). Landman (2001), a narrative psychologist, tells the true story of a fugitive who drove the getaway car in a heist that left a security guard dead. Decades later, this woman "obsessed with a desire to be punished, to seek expiation" (Franks 1994, p. 54) turned herself in to clueless authorities. Asked why she confessed, the now model citizen told her lawyer that "my strongest weapon against suicide is my contract with God . . ." (Franks 1994, p. 42).

In many societies, not only is supernatural punishment envisioned to fall directly upon the heads of the wicked, but is also believed to be sanguineous. Some supernatural agents are seen as unforgiving and merciless, inflicting lasting and far-reaching punishments across generations (Bering & Johnson 2005). Perhaps the worst punishment of all would be for one's biological relatives, especially offspring, to be cursed for the self's misdeeds. This is a particularly recurrent theme and is illustrated very clearly in the following brief passage on the Pagai from a Dutch missionary publication:

> A missionary once acted emphatically against various [superstitious] prohibitions in order to demonstrate their inefficacy. Actually this made a totally wrong impression on the people because they said: "The man knows perfectly well that he himself won't be punished but that the punishment will fall on his children." (Anonymous 1939, p. 9)

Recent laboratory findings suggest that there may be identifiable cognitive developmental milestones that promote the pan-cultural human tendency to see "signs" or hidden messages in natural events. In a study by Bering and Parker (in press), 3- to 9-year-old children were informed that an invisible agent (Princess Alice) would help them play a forced-choice game by "telling them, somehow, when they chose the wrong box," whereas a matched control group of children were not given this supernatural prime. On two *unexpected event* trials, as soon as the child chose a box, an experimenter triggered a simulated unexpected event in the laboratory (i.e., a light turning on/off; a picture falling), and children's response to these events (i.e., moving their hand to the opposite box) was coded. Thus, the study sought to determine the age at which children first begin to view natural events as being *about* their behaviors and as stemming from the mind of a communicative supernatural agent.

Results showed a significant interaction of age group by experimental condition. The only children to reliably move their hand to the opposite box in response to the unexpected events were the oldest children ($M = 7$ years, 4 months) who were primed with the invisible agent concept. Whereas 82% of these oldest children assigned to the experimental condition (and therefore told about Princess Alice) changed their response, only 18% of same-aged children in the control group (who were not told about Princess Alice) moved their hand to the opposite box after the unexpected event. For children's post-test verbal explanations, also, only the oldest children from the experimental group saw the unexpected events as being *referential* and *declarative* (e.g., "Princess Alice did it because I chose the wrong box"). In contrast, younger children ($M = 5$ years, 6 months) saw the event as being intentionally caused by the invisible agent (e.g., "Princess Alice did it because she wanted to"), whereas the preschoolers ($M = 4$ years, 1 month) did not invoke the invisible agent at all, but only physical causes for the event (e.g., "The picture fell because it wasn't sticking very well").

Although the cause of these age differences is controversial, these findings nevertheless demonstrate that the tendency to over-attribute intentions to the natural world is not simply a matter of hyperactive agency detection (e.g., Atran 2002; Atran & Norenzayan 2004; Barrett 2000; Guthrie 1993), but rather it *also* involves, at least in older children and adults, making inferences of communicative meaning within a social context. In this case, the specific supernatural agent concept (Princess Alice), which may be a rough analogue of culturally specific supernatural agents, appeared to map onto this inferential capacity for seeing signs in natural events. Furthermore, these *subjective* inferences gave way to *objective* behaviors, which is the primary currency of natural selection. An event such as a picture falling to the ground is not, in itself, a communicative event; it can become so only through the phenomenal properties of the child's mind ("what is the *meaning* of the picture crashing to the floor, precisely at this moment in time?").

5. Conclusion

In reviewing the available – though still very limited – evidence, there are good conceptual grounds to argue that natural selection may have set to work on specific human cognitive errors. These include simulation constraints leading to Type I errors in reasoning about the afterlife, teleo-functional errors leading to belief in the soul's intelligent design, and theory of mind errors fostering a belief that natural events were intentionally caused by supernatural agents. The resultant cognitive system created the functional illusion that the social behaviors of the self "mattered" outside of human relations. As a consequence it became morally tamed under the auspices of this existential rubric and therefore was less likely to engage in acts that, if publicly exposed and harmful to one's social reputation, seriously impaired genetic fitness.

The present article has also served to lay out some general future directions for investigators to more precisely explore the Darwinian mechanisms at the heart of the existential system outlined here. Such work can further reveal how the standard architecture of ancestral human minds was co-opted by natural selection to create the functional illusion of an intelligently designed, immortal soul that was under nearly unbreakable moralistic contract with the natural world.

ACKNOWLEDGMENTS

I am grateful to the following individuals for their careful reading and comments on earlier versions of this article: Joseph Bulbulia, Dominic Johnson, Shaun Nichols, Harvey Whitehouse, Nicholas Humphrey, and Paul Bloom. Also, I would like to thank several people who shared with me their thoughts and ideas on the topics discussed herein, including David Bjorklund, Todd Shackelford, and Juan Quiles. Jared Piazza pointed out to me Gay's epitaph quoted at the introduction of this article and Dan Gilbert informed me of Freud's remarks on the "afterlife" in *Thoughts for the Times on War and Death*. Finally, I wish to

acknowledge Jeff Gredlein, without whose assistance I would have been unable to complete this project.

NOTES

1. This is as in various physical resurrectionist beliefs, such as the Anabaptist doctrine of "soul-sleep," in which the soul is said to hibernate, or lie in wait, until it may reanimate the physically reconstituted body.

2. The simulation constraint hypothesis is indirectly supported by recent findings of egocentric social cognitive biases in adults (Epley et al. 2004). Epley and his colleagues found that participants' eye gaze preferentially moved to privileged visual space in response to an experimenter's ambiguous referential communication. For example, the command "move the bunny" elicited automatic eye gaze toward a stuffed bunny that could be seen by the participant, but which was occluded from the experimenter's perspective, over a chocolate Easter bunny to which both the participant and experimenter had visual access. The authors argue that these findings show that egocentrism is just as prevalent in adults as it is in young children. Adults, however, more rapidly correct their egocentrism to adjust for others' limited knowledge (e.g., by quickly shifting their gaze and moving the chocolate Easter bunny). If, as Epley et al. (2004) reason, individuals do become better with experience at making adjustments to correct for their initial egocentric views, but then rely on simulation to revise their social attributions, then even the best perspective-taking skills should falter when it comes to reasoning about dead agents' "perspective-less" minds. This is because any attempt at correcting for egocentrism by using simulation would still run up against simulation constraints (e.g., "does he *know* that he's dead?") and generate attributions of continued psychological functioning. Indeed, this is what is generally found.

3. The atrocities of the Holocaust forced many survivors to question God's "benevolent" intentions, apparently prompting some Jews to revise their theological views to accommodate the possibility that God is in fact morally corrupt. Nowhere is this theme more salient than in the semi-autobiographical chronicles of Holocaust survivor Elie Wiesel. In *Gates of the Forest*, Wiesel (1966, p. 197) writes:

> In a concentration camp, one evening after work, a rabbi called together three of his colleagues and convoked a special court. Standing with his head held high before them, he spoke as follows: "I intend to convict God of murder, for he is destroying his people and the law he gave to them . . . I have irrefutable proof in my hands. Judge without fear or sorrow or prejudice. Whatever you have to lose has long since been taken away." The trial proceeded in due legal form, with witnesses for both sides with pleas and deliberations. The unanimous verdict: "Guilty." . . . [But] after all, He had the last word. On the day of the trial, He turned the sentence against his judges and accusers. They, too, were taken off to the slaughter. And I tell you this: if their death has no meaning, then it's an insult, and if it does have a meaning, it's even more so.

4. In his *Bridge of San Luis Rey* (1927/1955), Thornton Wilder fictionalizes the sad tale of a collapsed bridge in eighteenth century Peru that brought five travelers to their deaths in the abyss below. In two chapters, one titled "Perhaps an Accident" and the other titled "Perhaps an Intention," Wilder describes how the resident monk, Brother Juniper, troubled by the seeming arbitrariness of this horrific event, embarks on a "scientific experiment" to reveal why God chose to end the lives of *these five* people rather than *some other five*, by collecting and analyzing the facts and details of each person's value in terms of goodness, piety, and usefulness. Alas, "the thing was more difficult than he had foreseen" and his quest for spiritual understanding went unresolved. In a case of life imitating art, 14 people lost their lives in 2001 when a runaway tugboat rammed two barges into an interstate bridge and caused about a dozen cars to collapse into the Arkansas River. One of the victims was a young army captain and father of four from California on his way home to Virginia. *The Oklahoman* newspaper reported that his commanding officer, echoing the thoughts of Brother Juniper, "pondered the odds of making a 2,929-mile drive and landing on a 500-foot stretch of bridge that, in the most bizarre of accidents, plummeted precisely as he crossed it. 'If [he] just stopped at a rest stop or stopped to get gas . . . There's just so many variables—and the timing.'" (Owen 2002).

Open Peer Commentary

Simulation constraints, afterlife beliefs, and common-sense dualism

Michael V. Antony

Department of Philosophy, University of Haifa, Haifa 31905, Israel.
antony@research.haifa.ac.il
http://philo.haifa.ac.il/faculty_pages/antony.htm

Abstract: Simulation constraints cannot help in explaining afterlife beliefs in general because belief in an afterlife is a precondition for running a simulation. Instead, an explanation may be found by examining more deeply our common-sense dualistic conception of the mind or soul.

Early on in his stimulating target article, Bering notes that the ability to conceive of an afterlife requires a dualistic conception of the relation between the conscious mind or soul and the body; and he is sympathetic (as I am also) to the idea that our common-sense concept of the mind/soul is dualistic, and in all likelihood innate. An important question for Bering is "how . . . we get from the common-sense dualism of infants to beliefs of the afterlife [. . .]" (target article, sect. 1, para. 4). And a major part of his answer is given by his "simulation constraint hypothesis," the idea that afterlife beliefs are explained by our attempts to mentally simulate "what it's like to be dead": putting ourselves "into the shoes" of dead agents, we are compelled to ascribe to them mental states.

While simulation constraints may help explain the specific types of mental states we project into the afterlife (as Bering argues), I do not think they can help explain why people believe in an afterlife in the first place. The point of a mental simulation, after all, is to generate conclusions about an agent's mental states or behaviors (with the type of simulation run depending on the types of mental states or behaviors about which one wishes to derive information). The cognitive mechanisms involved in planning simulations, accordingly, must assume the existence of a mind – namely, that mind into the nature of which one aims to gain insight through simulation. But this must hold for the afterlife case too: prior to simulating a dead agent's mind, it must be assumed there is a mind to simulate. But that *already* is to assume an afterlife. This mind/soul may be taken to be phenomenally rich, or relatively barren (experiencing "darkness," "nothingness," or what have you), but it must be taken to exist, at least implicitly. Notice that Bering seems to grant this in referring to "simulation strategies to derive information *about the minds* of dead agents" (sect. 2.1, para. 1, emphasis mine). It follows that nothing about a simulation itself can explain our belief in an afterlife, since some such belief or assumption is a precondition for the planning and running of any such simulation.

If that is right, how might afterlife beliefs be explained? I believe that the route from our common-sense dualism to