BEHAVIORAL AND BRAIN SCIENCES (2004) **27,** 713–770
*Printed in the United States of America*

# Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion

**Scott Atran**
*CNRS–Institut Jean Nicod, 75007 Paris, France; and Institute for Social Research–University of Michigan, Ann Arbor, MI 48106-1248*
**satran@umich.edu** **http://www.institutnicod.org**

**Ara Norenzayan**
*Department of Psychology, University of British Columbia, Vancouver, British Columbia V6T 1Z4 Canada*
**ara@psych.ubc.ca** **www.psych.ubc.ca/~ara**

**Abstract:** Religion is not an evolutionary adaptation per se, but a recurring cultural by-product of the complex evolutionary landscape that sets cognitive, emotional, and material conditions for ordinary human interactions. Religion exploits only ordinary cognitive processes to passionately display costly devotion to counterintuitive worlds governed by supernatural agents. The conceptual foundations of religion are intuitively given by task-specific panhuman cognitive domains, including folkmechanics, folkbiology, and folkpsychology. Core religious beliefs minimally violate ordinary notions about how the world is, with all of its inescapable problems, thus enabling people to imagine minimally impossible supernatural worlds that solve existential problems, including death and deception. Here the focus is on folkpsychology and agency. A key feature of the supernatural agent concepts common to all religions is the triggering of an "Innate Releasing Mechanism," or "agency detector," whose proper (naturally selected) domain encompasses animate objects relevant to hominid survival – such as predators, protectors, and prey – but which actually extends to moving dots on computer screens, voices in wind, and faces on clouds. Folkpsychology also crucially involves metarepresentation, which makes deception possible and threatens any social order. However, these same metacognitive capacities provide the hope and promise of open-ended solutions through representations of counterfactual supernatural worlds that cannot be logically or empirically verified or falsified. Because religious beliefs cannot be deductively or inductively validated, validation occurs only by ritually addressing the very emotions motivating religion. Cross-cultural experimental evidence encourages these claims.

**Keywords:** agency; death anxiety; evolution; folkpsychology; Maya; memory; metarepresentation; morality; religion; supernatural

## 1. Introduction

In every society,[1] there are

1. Widespread counterfactual and counterintuitive beliefs in supernatural agents (gods, ghosts, goblins, etc.)
2. Hard-to-fake public expressions of costly material commitments to supernatural agents, that is, offering and sacrifice (offerings of goods, property, time, life)
3. Mastering by supernatural agents of people's existential anxieties (death, deception, disease, catastrophe, pain, loneliness, injustice, want, loss)
4. Ritualized, rhythmic sensory coordination of (1), (2), and (3), that is, communion (congregation, intimate fellowship, etc.)

In all societies there is an evolutionary canalization and convergence of (1), (2), (3), and (4) that tends toward what we shall refer to as "religion"; that is, passionate communal displays of costly commitments to counterintuitive worlds governed by supernatural agents. Although these facets of religion emerge in all known cultures and animate the majority of individual human beings in the world, there are

SCOTT ATRAN is Director of Research (CNRS) at the Institute Jean Nicod in Paris and Adjunct Professor of Psychology at the University of Michigan. He received his Ph.D. in Anthropology from Columbia University. His research in Mesoamerica concerns universal and culture-specific aspects of biological categorization and reasoning and environmental decision making. He is currently interviewing Jihadists in the Middle East, exploring the limits of rational choice and the role of sacred values among suicide terrorists.

ARA NORENZAYAN is Assistant Professor of Psychology at the University of British Columbia. He was named an Early Career Scholar at the Peter Wall Institute of Advanced Studies in Vancouver, Canada in 2002–2003. The author of over 15 publications in the area of social and cultural psychology, he received his Ph.D. in Psychology from the University of Michigan in 1999. His interests include cognition across cultures, the psychological foundations of culture, and religious cognition.

© 2005 Cambridge University Press 0140-525X/04 $12.50

https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

considerable individual and cultural differences in the degree of religious commitment. The question as to the origin and nature of these intriguing and important differences we leave open.

This theoretical framework drives our program of research.[2] The framework is the subject of a recent book (Atran 2002a). Here, a more comprehensive set of experimental results and observations is introduced to support integration within an evolutionary perspective that envisions religion as a converging by-product of several cognitive and emotional mechanisms that evolved for mundane adaptive tasks (for somewhat similar, independently researched, views of religion as an emergent by-product of numerous domain-specific psychological mechanisms, see Boyer 2001; Kirkpatrick 1999b).

The current experiments suggestively support this long-term research program. We hope the findings will stimulate further tests and refinements to assess the empirical viability of this framework. The aim of this paper is to foster scientific dialogue between the fields of cultural anthropology, cognitive, developmental and social psychology, and evolutionary biology regarding a set of phenomena vital to most human life and all societies. The present article is mainly concerned with the first and third criteria of religion listed above. In this introductory section, we present in general terms the overall intellectual framework that interrelates all four criteria, discuss some obvious objections to these generalizations, and offer some caveats.

The criterion (1) of belief in the supernatural rules out commitment theories of religion as adequate, however insightful the latter may be. Such theories underplay or disregard cognitive structure and its causal role. Commitment theories attempt to explain the apparent altruism and emotional sacrifice of immediate self-interest accompanying religion in terms of long-term benefits to the individual (Alexander 1987; Irons 1996; Nesse 1999) or group (Boehm 1999; Wilson 2002) – benefits that supposedly contribute to genetic fitness or cultural survival. They do not account for the cognitive peculiarity of the culturally universal belief in beings who are imperceptible in principle, and who change the world via causes that are materially and logically inscrutable in principle. They cannot distinguish Marxism from monotheism, or secular ideologies from religious belief (Atran 2002a).

The criterion (2) of costly commitment rules out cognitive theories of religion as inadequate, however insightful they may be. Cognitive theories attempt to explain religious belief and practice as cultural manipulations of ordinary psychological processes of categorization, reasoning, and remembering (Andresen 2000; Atran & Sperber 1991; Barrett 2000; Boyer 1994; Lawson & McCauley 1990; Pyysiäinen & Anttonen 2002). They do not account for the emotional involvement that leads people to sacrifice to others what is dear to themselves, including labor, limb, and life. Such theories are often short on motive and are unable to distinguish Mickey Mouse from Moses, cartoon fantasy from religious belief (Atran 1998, p. 602; cf. Boyer 2000; Norenzayan & Atran 2004). They fail to tell us why, in general, the greater the sacrifice – as in Abraham offering up his beloved son – the more others trust in one's religious commitment (Kierkegaard 1843/1955).

We extend the idea (first suggested by Sperber 1975b) that religious thought and behavior can be explained as mediated by ordinary mental mechanisms, which can be scientifically studied regardless of whether religions are true or not true in a metaphysical sense. In this "mentalist" tradition, the focus so far has been on cognition and culture; that is, on how religious ideas are mentally constructed, transmitted across minds, and acquired developmentally. To be sure, there have been recent attempts by cognitive scientists studying religion to consider the role of emotion, and growing realization that religion cannot have a purely cognitive explanation that fails to take into account the social dilemmas motivating religious beliefs and practices (McCauley & Lawson 2002; Pyysiännen 2001; Whitehouse 2000). But there is still little analytic or empirical integration of (1) and (3).

Religions invoke supernatural agents (Horton 1967; Tylor 1871/1958) to deal with (3) emotionally eruptive existential anxieties (Malinowski 1922/1961), such as death and deception (Becker 1973; Feuerbach 1843/1972; Freud 1913/1990).[3] All religions, it appears, have "awe-inspiring, extraordinary manifestations of reality" (Lowie 1924, p. xvi). They generally have malevolent and predatory deities as well as more benevolent and protective ones. Supernatural agent concepts trigger our naturally selected agency-detection system, which is trip-wired to respond to fragmentary information, inciting perception of figures lurking in shadows and emotions of dread or awe (Guthrie 1993; cf. Hume 1757/1956). Granted, nondeistic "theologies," such as Buddhism and Taoism, doctrinally eschew personifying the supernatural or animating nature with supernatural causes. Nevertheless, common folk who espouse these faiths routinely entertain belief in an array of gods and spirits that behave counterintuitively in ways that are inscrutable to factual or logical reasoning.[4] Even Buddhist monks ritually ward off malevolent deities by invoking benevolent ones, and they perceive altered states of nature as awesome.[5]

Conceptions of the supernatural invariably involve the interruption or violation of universal cognitive principles that govern ordinary human perception and understanding of the everyday world. Consequently, religious beliefs and experiences cannot be reliably validated (or disconfirmed as false) through consistent logical deduction or consistent empirical induction. Validation occurs only by (4) collectively satisfying the emotions that motivate religion in the first place. Through a "collective effervescence" (Durkheim 1912/1995), communal rituals rhythmically coordinate emotional validation of, and commitment to, moral truths in worlds governed by supernatural agents. Rituals involve sequential, socially interactive movement and gesture, and formulaic utterances that synchronize affective states among group members in displays of cooperative commitment. Through the sensory pageantry of movement, sound, smell, touch, and sight, religious rituals affectively coordinate actors' minds and bodies into convergent expressions of public sentiment (Turner 1969) – a sort of *N*-person bonding that communicates moral consensus as sacred, transcending all reason and doubt (Rappaport 1999). Sensory pageantry also ensures the persistence and transmission of the religious beliefs and practices it infuses.

These four conditions do not constitute the necessary and sufficient features of "religion." Rather, they comprise a stipulative (working) framework that delimits a causally interconnected set of pancultural phenomena, which is the object of our study. One may choose to call phenomena that fall under this set of conditions "religion" or not; however, for our purposes the joint satisfaction of all four conditions


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

is what we mean by the term *religion.* Nevertheless, we offer this working framework as an adequate conceptualization that roughly corresponds to what most scholars consider religion. This framework is concerned with the pancultural foundations of religion; accordingly, our conceptualization is broad in scope. Surely, religions are manifested in culturally diverse ways and are shaped by local cultural contexts. Elsewhere, scholars have examined how the distinctive paths that religions take shape psychological tendencies (e.g., Shweder et al. 1997; Weber 1946). Our framework is not incompatible with these approaches. Indeed, it offers candidates for the psychological building blocks of religion, which then are culturally exploited in distinct but converging paths.

More critical are the many ethnographic reports which interpret that some people or some societies make no hard and fast distinction between (1), the natural and supernatural, or between (2), costly sacrifice and the social redistribution of material or social rewards; or that (3) religions are as anxiety-activating as they are anxiety-assuaging, or that (4) they are sometimes devoid of emotional ritual. In addition, (5) there is considerable psychological and sociological evidence for the health and well-being benefits of religion, which suggests that religion may be adaptive and not simply a by-product of evolutionary adaptations for other things. We address each of these objections next.

### 1.1. The natural versus the supernatural

We base our argument regarding the cognitive basis of religion on a growing number of converging cross-cultural experiments on "domain-specific cognition" emanating from developmental psychology, cognitive psychology, and anthropology. Such experiments indicate that virtually all (non-brain-damaged) human minds are endowed with core cognitive faculties for understanding the everyday world of readily perceptible substances and events (for overviews, see Hirschfeld & Gelman 1994; Pinker 1997; Sperber et al. 1995). The core faculties are activated by stimuli that fall into a few intuitive knowledge domains, including: folkmechanics (object boundaries and movements), folkbiology (biological species configurations and relationships), and folkpsychology (interactive agents and goal-directed behavior). Sometimes operation of the structural principles that govern the ordinary and "automatic" cognitive construction of these core domains are pointedly interrupted or violated, as in poetry and religion. In these instances, counterintuitions result that form the basis for construction of special sorts of counterfactual worlds, including the supernatural; for example, a world that includes self-propelled, perceiving, or thinking mineral substances (e.g., Maya *sastun,* crystal ball; Arab *tilsam* [talisman]) or beings that can pass through solid objects (angels, ghosts, ancestral spirits) (cf. Atran & Sperber 1991; Boyer 1994).

These core faculties generate many of the universal cognitions that allow cross-cultural communication and make anthropology possible at all. For example, even neonates assume that a naturally occurring rigid body cannot occupy the same space as another (unlike shadows), or follow discontinuous trajectories when moving through space (unlike fires), or change direction under its own self-propelling initiative (unlike animals), or causally effect the behavior of another object without physical contact (unlike people) (Spelke et al. 1995). When experimental conditions simulate violation of these universal assumptions, as in a magic trick, neonates show marked surprise (longer gaze, intense thumb sucking, etc.). Children initially expect shadows to behave like ordinary objects, and even adults remain uncertain as to how shadows move. This uncertainty often evokes the supernatural.

All known societies appear to partition local biodiversity into mutually exclusive species-like groupings (Atran 1990; Berlin 1992; Darwin 1859; Diamond 1966), and to initially identify nonhuman organisms according to these groupings rather than as individuals (unlike the immediate local identification of individual human faces and behaviors; Atran 1998; cf. Hirschfeld 1996). Individualized pets and taxonomic anomalies, such as monsters, become socially relevant and evocative because they are purposely divorced from the default state of "automatic" human cognition about the limited varieties of the readily perceptible world, that is, "intuitive ontology" (Atran 1989; Boyer 1997; cf. Sperber 1975b). This commonsense ontology is arguably generated by task-specific "habits of mind," which evolved selectively to deal with ancestrally recurrent "habits of the world" that were especially relevant to hominid (and in some cases, pre-hominid) survival, that is, inanimate substances, organic species, and persons.

What testable evidence there is indicates that, sometime after age three and except for severe autistics, most any person understands that most any other person can entertain perceptions, beliefs, and desires different from one's own, and that these different mental states differentially cause people's behaviors (Avis & Harris 1991; Baron-Cohen 1995; Knight et al. 2004; Wimmer & Perner 1983). Granted, there is experimental evidence for cultural variations in causal attribution of social behavior to personality traits versus social situations (Choi et al. 1999), and there are anecdotal interpretations of cultural behaviors indicating an inability to distinguish between true and false beliefs, or reality from desire (cf. Lévy-Bruhl 1923/1966; Lillard 1998). But contrary to the anecdotal evidence, experimental evidence suggests that children growing up in very different cultures soon develop similar understanding of core aspects of human behavior as a function of beliefs and desires (Avis & Harris 1991; Flavell et al. 1983). Furthermore, there is no generally accepted body of evidence indicating that our simian cousins can simultaneously keep in mind the thoughts of others, or, equivalently, entertain multiple possible and different worlds from which to select an appropriate course of action (Hauser 2000; Premack & Woodruff 1978; cf. Hare et al. 2001 for intriguing experiments suggesting rudimentary perspective taking in chimps). Without the ability to entertain multiple possible worlds, belief in the supernatural is inconceivable.

Within the emerging work on domain specificity there are controversies and doubts, as in any young and dynamic science. But the findings sketched above are widely replicated. Admittedly, there are alternative approaches to understanding cognition, such as connectionism, artificial intelligence, and phenomenology. Using any of these other approaches to model religion would no doubt present a different picture than the one we offer. We leave it to others to work the alternatives.

### 1.2. Costly sacrifice versus redistribution

One evolutionary problem with religion is explaining how and why biologically unrelated individuals come to sacrifice



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

their own immediate material interests to form genetically incoherent relationships under an imagined permanent and immaterial authority. Altruism occurs when an organism's behavior diminishes its own fitness and enhances the fitness of some other organism or organisms. Fitness is a measure of an organism's reproductive success. The sacrifice of an organism for its relatives – a mother for her children, a brother for his siblings, an ant for its colony, a bee for its hive – lowers an organism's individual fitness (also called "classical" or "Darwinian" fitness) because it compromises the individual's ability to bear and raise offspring. Nevertheless, such kin altruism may also enhance the individual's "inclusive fitness" by allowing surviving relatives to pass on many of the individual's genes to future generations (Hamilton 1964). But what motivates the sort of *non-kin* cooperation characteristic of human religious commitment?

Unlike other primate groups, hominid groups grew to sizes (Dunbar 1996) that could not function exclusively on the basis of kin selection (commitment falls off precipitously as genetic distance increases between individuals) or direct reciprocity (ability to directly monitor trustworthiness in reciprocation decreases rapidly as the number of transactions multiply). Larger groups of individuals outcompete smaller groups in love and war (Axelrod 1984). A plausible hypothesis, then, is that the mechanisms for successful promotion of indirect reciprocity – including both religious and nonreligious behaviors – were naturally selected in response to the environmental problem-context of spiraling social rivalry among fellow conspecifics, or "runaway social competition" (Alexander 1989). As "fictive kin" (Nesse 1999), members of religious groups perform and profit from many tasks that they could not do alone, one by one, or only with family. Thus, "Among the Hebrews and Phoenicians . . . the worshipper is called brother (that is, kinsman or sister of the god)" (Robertson Smith 1891/1972, p. 44, note 2). "Brotherhood" is also the common term applied today among the Christian faithful and to the fraternity (*ikhwan*) of Islam.

Indirect reciprocity occurs when individual *X* knows that individual *Y* cooperates with others, and this knowledge favors *X* cooperating with *Y*. Consider a population whose individuals have the option to cooperate or not. Suppose individual *X* randomly meets individual *Y*. If *Y* has a reputation for cooperation, and if *X* cooperates with *Y*, then *X*'s reputation likely increases. If *X* does not cooperate with *Y*, then *X*'s reputation likely decreases (see Nowak & Sigmund 1998 for various simulations). The basic idea is to help those who are known to help others. Reputation for religious belief is almost always reckoned as sincere social commitment, and such reputation is invariably linked to costly and hard-to-fake expressions of material sacrifice or concern that goes beyond any apparent self-interest.

Although calculations of economic or political utility often influence religious practices (Stark 2000), to conclude that that's all there is to religious commitment and sacrifice is unwarranted. In religious offerings, there is *usually a nonrecuperable cost* involved both in the selection of the item offered and in the ceremony itself. Thus, for the Nuer of Sudan, substituting a highly valued item (cow) by one that is less valued (fowl or vegetable) is allowable only to a point, after which "a religious accounting might reveal that the spirits and ghosts were expecting a long overdue proper sacrifice, because accounts were out of balance, so to speak" (Evans-Pritchard 1940, p. 26). Religious sacrifice usually costs something for the persons on whose behalf the offering is made. That is why "sacrifice of wild animals which can be regarded as the free gift of nature is rarely allowable or efficient" (Robertson Smith 1894, p. 466). In many cases, the first or best products of one's livelihood goes to the gods, as with the first fruits of the Hebrews or the most perfect maize kernels of the Maya. Most, if not all, societies specify obligatory circumstances under which religious sacrifice *must* be performed, *regardless* of economic considerations. Reviewing the anthropological literature, Raymond Firth (1963, p. 16) surmises, "In all such cases the regular religious need to establish communication with god or with the spirit world . . . would seem to be pressing and primary. 'Afford it or not.'"

In sum, religious sacrifice generally runs counter to calculations of immediate utility, such that future promises are not discounted in favor of present rewards. In some cases, sacrifice is extreme. Although such cases tend to be rare, they are often held by society as religiously ideal: for example, sacrificing one's own life or nearest kin. Researchers sometimes take such cases as prima facie evidence of "true" (non-kin) social altruism (Kuper 1996; Rappaport 1999) or group selection, wherein individual fitness decreases so that overall group fitness can increase (relative to the overall fitness of other, competing groups) (Sober & Wilson 1998; Wilson 2002). But this may be an illusion.

A telling example is contemporary suicide terrorism (Atran 2003a). Through indoctrination of recruits into relatively small and closeted cells – emotionally tight-knit "brotherhoods" – terror organizations create a "family" of cell mates who are just as willing to sacrifice for one another as a mother for her children. Consider the "Oath to Jihad" taken by recruits to *Harkat al-Ansar,* a Pakistan-based ally of Al-Qaida, which affirms that by their sacrifice, they will help secure the future of their "family" of fictive kin: "Each [martyr] has a special place – among them are brothers, just as there are sons and those even more dear." These culturally contrived cell loyalties mimic and (at least temporarily) override genetically based fidelities to family kin while securing belief in sacrifice to a larger group cause. The mechanism of manipulation resembles the one used by our own army to train soldiers in small groups of committed buddies who acquire willingness to sacrifice for one another, and, derivatively, for glory and country (motherland, fatherland). In the case of religiously inspired suicide terrorism, these sentiments are purposely manipulated by organizational leaders, recruiters, and trainers to the advantage of the manipulating elites rather than the individual (much as the fast-food or soft-drink industries manipulate innate desires for naturally scarce commodities like fatty foods and sugar to ends that reduce personal fitness but benefit the manipulating institution). No "group selection" is involved, only cognitive and emotional manipulation of some individuals by others.

### 1.3. Relieving versus provoking anxieties

Often the naturally eruptive anxieties that bring on the supernatural are artificially (purposely) excited, then assuaged (Durkheim 1912/1995). It might seem, then, that the problem of religion's ability to neutralize suffering is akin to the wag about the salesman who throws dirt on the rug in order to demonstrate the vacuum cleaner's ability to remove it. Consider initiation rituals that involve "rites of terror" (Whitehouse 1996), as among Native American Cheyenne and Arapaho (Lowie 1924), Walbiri (Meggitt 1965) and other aboriginals of the Central Australian


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

Desert (Spencer & Gillen 1904), Mountain Ok Baktaman (Barth 1975) and Ilahita Arapesh of Highland Papua New Guinea (Tuzin 1982), and Candombolé Nagô sects of African-Brazilian Bahia (Carneiro 1940; Omari 1994). These arouse existential anxieties by culturally mimicking and manipulating seemingly capricious and uncontrollable situations that naturally provoke them: terror and risk of death from unidentifiable sources, the menace of infirmity and starvation through physical ordeal and deprivation, the injustice of whimsical oppression, sudden isolation, and loneliness. Often initiates temporarily manifest behaviors and cognitions associated with persons clinically diagnosed as suffering abuse, stress, or trauma, including re-experiencing the events (nightmares, intrusive memories, flashbacks), avoidance (amnesia of the event, refusal to talk or think about it), and hyperarousal (startle response, fitful sleep, poor concentration) (cf. Newport & Nemeroff 2000).

Still, there are important differences between such initiations and stress syndromes (e.g., posttraumatic stress disorder). Stress sufferers who permanently lose memory and undergo reduced immune response often suffer from chronic stress and lack of effective social support (Dhabar & McEwen 1999; Khansari et al. 1990). By contrast, even the most severe and emotionally aversive religious initiations end in positive exhibitions of social acceptance:

> Boys and girls are made to recognize members of The People [Navajo] and are introduced to full participation in ceremonial life. . . . The first boy is led out beside the fire. The figure in the white mask makes a mark on each shoulder with sacred cornmeal. . . . Then, using a different falsetto cry, the black-masked figure lightly strikes . . . other places on the body, and the one who uses the reeds varies the time interval between touching the boy and uttering his cry, so its unexpectedness causes the boy to start convulsively. . . . Then the one who wore the black mask places it over the face of each child in turn. . . . All the children are told to look up and always remember the Holy People. The reversal of the masks is a very intelligent psychological act, for it allows the child to see that the dread figure is actually someone he knows, or at least a human being, and thus the ritual is robbed of some of its terror. . . . The ceremony closes with the admonition to each child not to betray to uninitiates what he has seen. (Kluckholn & Leighton 1946/1974, p. 207–208; cf. Turnbull 1962, p. 225)

Through the stress that these exaggerated sensual displays induce, rites of passage furnish emotionally costly and memorable – but ultimately satisfying – commitments to the group and its supernatural agents.

In brief, these life rehearsals incite the very emotions and existential anxieties that motivate religious beliefs and quests for deliverance. Then, by assuaging and resolving the ensuing distress, successful completion of the ritual performance authenticates the religious thoughts and actions. This confirms the efficacy of religious belief and ritual performance in fusing cosmos to culture by overcoming the dreads and uncertainties of both spontaneously occurring natural events and the manipulated happenings of the social world.

### 1.4. Emotional ritual

Although there is wide variation in the degree of sensory pageantry associated with religious rituals (McCauley & Lawson 2002; Whitehouse 2000), religious rituals habitually – perhaps invariably – include displays of social hierarchy and submission typical of primates and other social mammals (outstretched limbs, baring throat and chest or genitals, genuflection, bowing, prostration, etc.). Even priests and kings must convincingly show sincere submission to higher supernatural authority lest their own authority be doubted (Aristotle 1958; Burkert 1996; cf. Watanabee & Smuts 1999).

Most often, religious rituals involve repeated, generally voluntary, and usually reversible states of emotional communion in the context of formulaic social ceremonies. Here, supernatural agents, through their surrogates and instruments, manifest themselves in people's affections. The ceremonies repetitively occur to make highly improbable, and therefore socially unmistakable, displays of mutual commitment. Within the congregation's coordinated bodily rhythms (chanting, swaying, tracking, etc.), in conjunction with submission displays, individuals show that they feel themselves identifying with, and giving over part of their being to, the intensely felt existential yearnings of others. This demonstration, in turn, conveys the *intention* or *promise* of self-sacrifice by and towards others (charity, care, defense, support, etc.), without any specific person or situation necessarily in mind.

Collective religious ritual always seems to involve ancestrally primitive communicative forms that Tinbergen calls "ritualized social releasers" (1951, p. 191–92). Social releasers exhibit sense-evident properties, "either of shape, or colour, or special movements, or sound, or scents," which readily elicit a well-timed and well-oriented cooperative response in a conspecific: for mating, parenting, fighting, defense, food gathering, and the like. But humans appear to be the only animals that spontaneously engage in creative, rhythmic bodily coordination to enhance cooperation. Unlike, say, avian mating calls or flight formations, human music or body dance, which are omnipresent in worship, can be arbitrarily and creatively recomposed.

A key feature of the creativity of human worship is use of music in social ritual. Even the Taliban, who prohibited nearly all public displays of sensory stimulation, promoted a cappella religious chants. Nearer to home, in a survey of persons who reported a religious experience (Greeley 1975), music emerges as the single most important elicitor of the experience. Listeners as young as 3 years old reliably associate basic or primary emotions to musical structures, such as happiness, sadness, fear, and anger (Trainor & Trehub 1992; cf. Panksepp 1993, Schmidt & Trainor 2002).

Much of the intimate connection between music and religion remains a puzzle. One possible account sees music as an invitation to interpersonal relationships, creating emotional bonds among people, through the "attunement" of somatic states – much as the rocking and cooing behavior of mother and infant attunes the parental bond (Stern 1985). This is especially apparent in a *call-response* format, as in Yoruba dances and Hebrew services. Moreover, in religious contexts, music is frequently experienced as authorless, like the sacred texts that often accompany it.[6] The pretonal religious music of small-scale societies usually has its mythic beginnings in the origins of the world, which invites audiences to share in a sense of timeless intimacy. For the Catholic Church, Gregorian chants were taught to men by birds sent from heaven. Even Bach, Mozart, and Beethoven were but vehicles of The Divine's call to communion.

### 1.5. "Mind-blind" functionalism: Sociobiology, group selection, and memetics

Finally, our account opposes other evolutionary approaches to religion and culture, including much sociobiology (Har-



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

ris 1974; Wilson 1978), group-selection theory (Boehm 1999; Sober & Wilson 1998), and memetics (Dawkins 1976; Dennett 1997). These alternatives are "mind-blind" to the cognitive constraints on religious beliefs and practices, viewing religion and culture as bundles of functionally integrated, fitness-bearing traits: for example, packages of environment-induced rituals (the material infrastructure underlying ideational superstructure), machinelike patternings of collective norms (worldviews), or partnerships of invasive and authorless ideas (memeplexes).

Proponents of these alternatives do not deny that minds have causally "proximate" roles in generating religious behaviors – as they may in generating economic behaviors – or that cognition may form part of some "ultimate" explanation of religion. Nevertheless, a common claim is that a meaningful causal account of such behaviors requires initial focus on measurable relationships between putative fitness-motivating factors in religious behaviors and ostensible fitness consequences (Dennett 1995, pp. 358–59; Sober & Wilson 1998, pp. 182, 193; cf. Lumsden & E.O.Wilson 1981): for example, between individuals needing protein in animal-poor environments and ritual human sacrifice (Harris 1974; Wilson 1978), between ideas endeavoring to propagate themselves and proselytizing for altruism (Blackmore 1999; Lynch 1996), or between groups competing for survival and Judaism's alleged cultural and genetic separatism (MacDonald 1998; Wilson 2002). These arguments are presented through selective use of anecdotal evidence, rather than being reliably tested and demonstrated.

Thus, despite sociobiological claims to the canons of "scientific materialism," the causal account that is supposed to produce religious practices (e.g., Aztec cannabilistic sacrifice) from their ostensible material functions (e.g., compensating for lack of large game as sources of protein in Mesoamerica) is wholly mysterious (e.g., How does eating someone generate the idea or formation of a pyramid or priest?). Moreover, similar practices often arise or endure independently or regardless of material need. For example, the African Azande said they just preferred the taste of human meat (Evans-Pritchard 1960), and game was abundant for Mesoamerica's Lowland Maya, who also practiced human sacrifice (de Landa 1566/1985).

It is also notoriously difficult to establish measurable criteria by which whole cultures/societies or worldviews/memeplexes can have fitness consequences.[7] Functional accounts are often synthetic abstractions: for example, a lone anthropologist's normative digest of some culture that in reality has no clear boundaries and no systematically identifiable structural functions. Indeed, most reported "norms" are too semantically open-ended to have specific contents, such as the Ten Commandments: even members of the same church congregation fail to provide interpretations of the Ten Commandments that other congregation members consistently recognize as being interpretations of the Ten Commandments (Atran 2001b). There are no "replicating" or even definite or definable cultural units for natural selection and vertical (transgenerational) or horizontal (contemporaneous) transmission (e.g., memes can be anything from a gender marker to partial tune, cell phone, cooking recipe, political philosophy, etc.). These facts render implausible all attempts to explain religions (or cultures with a religious element) as discrete or integrated functional systems (for reviews and analyses of specific arguments, see Atran 2001b; 2002a; 2003b; in press).

All human societies pay a price for religion's material, emotional, and cognitive commitments to unintuitive, factually impossible worlds. Functional evolutionary ("adaptationist") arguments for religion often try to offset its clear functional disadvantages with greater functional benefits. There are many different and contrary explanations as to why religion exists in terms of beneficial functions served. These include functions of social (bolstering group solidarity, group competition), economic (sustaining public goods, surplus production), political (mass opiate, rebellion's stimulant), intellectual (e.g., explain mysteries, encourage credulity), health and well-being (increase life expectancy, accept death), and emotional (terrorizing, allaying anxiety) utility. Many of these functions have obtained in one cultural context or another; yet all also have been true of cultural phenomena besides religion.

Such descriptions of religion are not wrong; however, none of these accounts provides explanatory insight into cognitive selection factors responsible for the ease of acquisition of religious concepts by children, or for the facility with which religious practices and beliefs are transmitted across individuals. They have little to say about which beliefs and practices – all things being equal – are most apt to survive within a culture, most likely to recur in different cultures, and most disposed to cultural variation and elaboration. None predicts the cognitive peculiarities of religion, such as

Why do *agent* concepts predominate in religion?

Why are *supernatural-agent* concepts culturally universal?

Why are *some* supernatural agent concepts *inherently better* candidates for cultural selection than others?

Why is it necessary, and how it is possible, to *validate* belief in supernatural agent concepts that are logically and factually inscrutable?

How is it possible to prevent people from deciding that the existing moral order is simply wrong or *arbitrary* and from *defecting* from the social consensus through denial, dismissal, or *deception*?

Our argument does not entail that religious beliefs and practices cannot perform social functions, or that the successful performance of such functions does not contribute to the survival and spread of religious traditions. Indeed, there is substantial evidence that religious beliefs and practices often alleviate potentially dysfunctional stress and anxiety (Ben-Amos 1994; Worthington et al. 1996) and maintain social cohesion in the face of real or perceived conflict (Allport 1956; Pyszczynski et al. 1999). It does imply that social functions are not evolutionarily responsible for the cognitive structure and cultural recurrence of religion. This article addresses these and related issues with cross-cultural experiments and observations.

## 2. The supernatural agent: Hair-triggered folkpsychology

Religions invariably center on *supernatural agent concepts*, such as gods, goblins, angels, ancestor spirits, jinns. In this section, we concentrate on the concept of *agency*, a central player in what cognitive and developmental psychologists refer to as "folkpsychology" and the "Theory of Mind" (ToM). *Agency*, we speculate, evolved hair-triggers in humans to respond "automatically" under conditions of un-


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

certainty to potential threats (and opportunities) by intelligent predators (and protectors). From this perspective, *agency* is a sort of "Innate Releasing Mechanism" (Tinbergen 1951) whose proper evolutionary domain encompasses animate objects, but which inadvertently extends to moving dots on computer screens, voices in the wind, faces in the clouds, and virtually any complex design or uncertain circumstance of unknown origin This insight into the supernatural as the by-product of a hair-triggered agency detector was first elaborated by Guthrie (Guthrie 1993; cf. Hume 1756/1957). We further ground it in the emerging theory of folkpsychology.

A number of experiments show that children and adults spontaneously interpret the contingent movements of dots and geometrical forms on a screen as interacting agents who have distinct goals and internal motivations for reaching those goals (Bloom & Veres 1999; Csibra et al. 1999; Heider & Simmel 1944; Premack & Premack 1995).[8] Such a biologically prepared, or *modular,* processing program would provide a rapid and economical reaction to a wide – but not unlimited – range of stimuli that would have been statistically associated with the presence of agents in ancestral environments. Mistakes, or "false positives," would usually carry little cost, whereas a true response could provide the margin of survival (Geary & Huffman 2002; Seligman 1971).

Our brains, it seems, are trip-wired to spot lurkers (and to seek protectors) where conditions of uncertainty prevail (when startled; at night; in unfamiliar environments; during sudden catastrophe; or in the face of solitude, illness, or prospects of death, etc.). Plausibly, the most dangerous and deceptive predator for the genus *Homo* since the Late Pleistocene has been *Homo* itself, which may have engaged in a spiraling behavioral and cognitive arms race of individual and group conflicts (Alexander 1989). Given the constant menace of enemies within and without, concealment, deception, and the ability to generate and recognize false beliefs in others would favor survival. In potentially dangerous or uncertain circumstances, it would be best to anticipate and fear the worst of all likely possibilities: presence of a deviously intelligent predator. How else could humans have managed to constitute and survive such deadly competitive groups as the Iatmul head-hunters of New Guinea (Bateson 1958) or the Nāga of Nagaland (north-eastern India)?

> All the Nāga tribes are, on occasion, head-hunters, and shrink from no treachery in securing these ghastly trophies. Any head counts, be it that of a man, woman, or child, and entitles the man who takes it to wear certain ornaments according to the custom of the tribe or village. Most heads are taken . . . not in a fair fight, but by methods most treacherous. As common a method as any was for a man to lurk about the water Ghāt of a hostile village, and kill the first woman or child who came to draw water. . . . Every tribe, almost every village is at war with its neighbour, and no Nāga of these parts dare leave the territory of his tribe without the probability that his life will be the penalty. (Crooke 1907, p. 41–43)

Throughout the world, societies cast their enemies as physically or mentally warped supernatural beings, or at least in league with the supernatural. Originally, *nāga* "applied to dreaded mountain tribes, and [was] subsequently used to designate monsters generally" (Werner 1932/1961, p. 284). The dragons of ancient India (*nāga*) and their Chinese derivatives (*lung*) are often depicted as creatures half human and half animal who emerge from the clouds to wreak havoc on humankind. Similarly, serpent-like devils and demons are culturally ubiquitous (Munkur 1983), perhaps evoking and addressing a primal fear shared by our primate line (Mineka et al. 1984).[9]

From an evolutionary perspective, it is better to be safe than sorry regarding the detection of agency under conditions of uncertainty. This cognitive proclivity would favor emergence of malevolent deities in all cultures, just as the countervailing Darwinian propensity to attach to protective caregivers would favor the apparition of benevolent deities. Thus, for the Carajá Indians of Central Brazil, intimidating or unsure regions of the local ecology are religiously avoided:

> The earth and underworld are inhabited by supernaturals. . . . There are two kinds. Many are amiable and beautiful beings who have friendly relations with humans. . . . The others are ugly and dangerous monsters who cannot be placated. Their woods are avoided and nobody fishes in their pools. (Lipkind 1940, p. 249)

Nearly identical descriptions of supernaturals can be found in ethnographic reports throughout the Americas, Africa, Eurasia, and Oceania (Atran 2002a).

In addition, humans *conceptually create* information to mimic and manipulate conditions in ancestral environments that originally produced and triggered our evolved cognitive and emotional dispositions (Sperber 1996). Humans habitually "fool" their own innate releasing programs, as when people become sexually aroused by make-up (which artificially highlights sexually appealing characteristics), fabricated perfumes, or undulating lines drawn on paper or dots arranged on a computer screen, that is, pornographic pictures.[10] Indeed, much of human culture – for better or worse – can be arguably attributed to focused stimulations and manipulations of our species' innate proclivities.

These manipulations can activate and play upon several different cognitive and emotional faculties at once. Thus, masks employ stimuli that trigger our innate, hyperactive facial-recognition schema. Masks also employ stimuli that activate, amplify, and confound emotions by highlighting, exaggerating, or combining certain facial expressions. Moreover, like two-dimensional drawings of the Nekker cube for which there is no stable three-dimensional interpretation, masks can produce feelings of unresolved anxiety or "uncanniness." In many religious ceremonies, for example, as a mask rotates away (e.g., clockwise) from an onlooker, who now gazes on the mask's hollow back, the onlooker perceives a three-dimensional face emerging in the other direction (counterclockwise) from inside the back of the mask (cf. Dawkins 1998). Such manipulations can serve cultural ends far removed from the ancestral adaptive tasks that originally gave rise to those cognitive and emotional faculties triggered, although manipulations for religion often centrally involve the collective engagement of existential desires (e.g., wanting security) and anxieties (e.g., fearing death).

Recently, numbers of devout American Catholics eyed the image of Mother Theresa in a cinnamon bun sold at a shop in Tennessee. Latinos in Houston prayed before a vision of the Virgin of Guadalupe, whereas Anglos saw only the dried remnants of melted ice cream on a pavement. Cuban exiles in Miami spotted the Virgin in windows, curtains, and television afterimages as long as there was hope



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

of keeping young Elian Gonzalez from returning to godless Cuba. And on the day of the World Trade Center bombing, newspapers showed photos of smoke billowing from one of the towers that "seems to bring into focus the face of the Evil One, complete with beard and horns and malignant expression, symbolizing to many the hideous nature of the deed that wreaked horror and terror upon an unsuspecting city" ("Bedeviling: Did Satan Rear His Ugly Face?," *Philadelphia Daily News,* 14 Sept. 2001). In all these cases, there is culturally conditioned emotional priming in anticipation of agency. This priming, in turn, amplifies the information value of otherwise doubtful, poor, and fragmentary agency-relevant stimuli. This enables the stimuli (e.g., cloud formations, pastry, ice cream conformations) to achieve the minimal threshold for triggering hyperactive facial-recognition and body-movement recognition schemata that humans possess.

In sum, supernatural agents are readily conjured up because natural selection has trip-wired cognitive schema for agency detection in the face of uncertainty. Uncertainty is omnipresent; so, too, is the hair-triggering of an agency-detection mechanism that readily promotes supernatural interpretation and is susceptible to various forms of cultural manipulation. Cultural manipulation of this modular mechanism and priming facilitate and direct the process. Because the phenomena created readily activate intuitively given modular processes, they are more likely to survive transmission from mind to mind under a wide range of different environments and learning conditions than entities and information that are harder to process (Atran 1998; 2001b). As a result, they are more likely to become enduring aspects of human cultures, such as belief in the supernatural.

## 3. Counterintuitive worlds

In this section we unpack the idea of the supernatural as a counterintuitive world that is not merely counterfactual in the sense of physically implausible or nonexistent. Rather, the supernatural literally lacks truth conditions. A counterintuitive thought or statement can take the surface form of a proposition (e.g., "Omnipotence [i.e., God] is insubstantial"), but the structure of human semantics is such that no specific meaning can be given to the expression and no specific inferences generated from it (or, equivalently, any and all meanings and inferences can be attached to the expression). The meanings and inferences associated with the subject (omnipotence = physical power) of a counterintuitive expression contradict those associated with the predicate (insubstantial = lack of physical substance), as in the expressions "the bachelor is married" or "the deceased is alive."[11]

All the world's cultures have religious myths that are attention-arresting because they are counterintuitive.[12] Still, people in all cultures also recognize that such beliefs are counterintuitive, whether or not they are religious believers (Atran 1996). Among Christian communities all over the world, for example, Catholics and non-Catholics alike are unquestionably aware of the difference between Christ's body and ordinary wafers, or between Christ's blood and ordinary wine. Likewise, Native American Cowlitz are well aware of the difference between the deity Coyote and everyday coyotes, or between Old Man Wild Cherry Bark and ordinary wild cherry bark (Jacobs 1934, p. 126–33).

Religious beliefs are counterintuitive because they violate what studies in cognitive anthropology and developmental psychology indicate are universal expectations about the world's everyday structure, including such basic categories of *intuitive ontology* (i.e., the ordinary ontology of the everyday world that is built into the language learner's semantic system) as *person, animal, plant,* and *substance* (Atran 1989). They are generally inconsistent with fact-based knowledge, though not randomly. Beliefs about invisible creatures who transform themselves at will or who perceive events that are distant in time or space flatly contradict factual assumptions about physical, biological, and psychological phenomena (Atran & Sperber 1991). Consequently, these beliefs more likely will be retained and transmitted in a population than random departures from common sense, and thus become part of the group's culture. Insofar as category violations shake basic notions of ontology, they are attention-arresting, hence memorable. But only if the resultant impossible worlds remain bridged to the everyday world can information be stored, evoked, and transmitted.

As a result, religious concepts need little in the way of overt cultural representation or instruction to be learned and transmitted. A few fragmentary narrative descriptions or episodes suffice to mobilize an enormously rich network of implicit background beliefs (Boyer 1994). For instance, if God is explicitly described as being jealous and able to move mountains, He is therefore implicitly known to have other emotions, such as anger and joy, and other powers, such as the ability to see and touch mountains or to lift and sight most anything smaller than a mountain, such as a person, pot, pig, or pea.

Invocation of supernatural agents implicates two cognitive aspects of religious belief: (1) activation of naturally selected conceptual modules, and (2) failed assignment to universal categories of ordinary ontology. Conceptual modules are activated by stimuli that fall into a few intuitive knowledge domains, including: folkmechanics (object boundaries and movements), folkbiology (species configurations and relationships), and folkpsychology (interactive and goal-directed behavior). Ordinary ontological categories are generated by further, more specific activation of conceptual modules. Among the universal categories of ordinary ontology are: *person, animal, plant, substance.*[13]

To give an example, sudden movement of an object stirred by the wind may trigger the agent-detection system that operates over the domain of folkpsychology, and a ghost may be invoked to interpret this possibly purposeful event. In normal circumstances, a sudden movement of wind might activate cognitive processing for agents, but would soon deactivate upon further analysis ("it's only the wind"). But in the case of (bodiless) supernatural agents, the object-boundary detectors that operate over the domain of folkmechanics, and which are required to identify the agent, cannot be activated. The same cognitive conditions operate when supernatural beings and events, like ghosts or gods, are evoked in religious ceremonies, whether or not there is any actual triggering event (e.g., a sudden movement of unknown origin or other uncertain happening). In such cases, assignment to the *person* or *animal* category cannot be completed because ghosts and gods have


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

counterintuitive properties (e.g., movements and emotions without physical bodies). This results in a potentially endless, open-textured evocation of possible meanings and inferences to interpret the event. However, the process can be provisionally stopped, and the semantic content somewhat specified, in a given context (e.g., a Sunday sermon that fixes interpretation of a Biblical passage on some particular community event in the preceding week).

Ordinary ontological categories always involve more specific processing over the folkmechanics domain (nonliving objects and events).

Only *substance* involves further processing that is exclusive to folkmechanics.

*Plant* involves additional processing over the folkbiological domain (every organism is assigned to one and only one folk species).

*Animal* involves supplemental processing over the domains of folkbiology (every animal is assigned uniquely to a folk species) as well as folkpsychology (animal behavior is scrutinized as indicating predator or prey, and possibly friend or foe).

*Person* involves more specific processing over the folkpsychological domain (human behavior is scrutinized as indicating friend or foe, and possibly predator or prey) and the folkbiological domain (essentialized group assignments, like race and ethnicity).

The relationship between conceptual modules and ontological categories is represented as a matrix in Table 1. Changing the intuitive relationship expressed in any cell generates what Boyer (2000) calls a "minimal counterintuition" (cf. Barrett 2000). For example, switching the cell (− folkpsychology, *substance*) to (+ folkpsychology, *substance*) yields a thinking talisman, whereas switching (+ folkpsychology, *person*) to (− folkpsychology, *person*) yields an unthinking zombie.

These are general, but not exclusive, conditions on supernatural beings and events. Intervening perceptual, contextual or psycho-thematic factors, however, can change the odds. Thus, certain natural substances – mountains, seas, clouds, sun, moon, planets – are associated with perceptions of great size or distance, and with conceptions of grandeur and continuous or recurring duration. They are, as Freud surmised, psychologically privileged objects for focusing the thoughts and emotions evoked by existential anxieties like death and eternity. Imaginary or actual violation of fundamental social norms also readily lends itself to religious interpretation (e.g., ritual incest, fratricide, status reversal).

Finally, supernatural agent concepts tend to be emotionally powerful because they trigger evolutionary survival templates. This also makes them attention-arresting and memorable. For example, an all-knowing bloodthirsty deity is a better candidate for cultural survival than a do-nothing deity, however omniscient. In the next section, we address some of the cognitive processes that contribute to the cultural survival of supernatural beliefs.

## 4. Cultural survival: A memory experiment

Many factors are important in determining the extent to which ideas achieve a cultural level of distribution. Some are ecological, including the rate of prior exposure to an idea in a population; physical, as well as social, facilitators and barriers to communication and imitation; and institutional structures that reinforce or suppress an idea. Other factors are psychological, including the cognitive and emotional ease with which an idea can be accommodated, represented, and remembered; the intrinsic interest that it evokes in people so that it is processed and rehearsed; and motivation and facility to communicate the idea to others.

One complex of psychological factors concerns the apparent sensitivity to religious ideas in young children. Studies of American and European children indicate that most children through grade 1 (ages 6–7) think that God is present everywhere, can hear prayers and see everything, and is near when one feels troubled or happy. This lends credence to the Jesuits' mantra of "Give me a child till the age of seven and I'll give you a Believer for life." Sentiments about God's pervasiveness in life seem to degrade with age unless institutionally supported, and God's presence and guidance become associated more with danger and difficulties (Goldman 1964; Tamminen 1994; Thun 1963).

Of all cognitive factors, however, mnemonic power may be the single most important one at any age (Sperber 1996). In oral traditions that characterize most of human cultures throughout history, an idea that is not memorable cannot be transmitted and cannot achieve cultural success (Rubin 1995). Moreover, even if two ideas pass a minimal test of memorability, a *more* memorable idea has a transmission advantage over a less memorable one (all else being equal). This advantage, even if small at the start, accumulates from generation to generation of transmission leading to massive differences in cultural success at the end.

One of the earliest accounts of the effects of memorabil-

Table 1. *Mundane relations between naturally selected conceptual domains and universal categories of ordinary ontology (changing the relation in any one cell (+ to −, or − to +) yields a minimal, supernatural counterintuition)*

| ONTOLOGICAL CATEGORIES | Conceptual Domains (and associated properties) | | | | |
|---|---|---|---|---|---|
| | Folkmechanics | Folkbiology | | Folkpsychology | |
| | (Inert) | (Vegetative) | (Animate) | (Psycho-Physical, e.g., hunger, thirst, etc.) | (Epistemic, e.g., believe, know, etc.) |
| PERSON | + | + | + | + | + |
| ANIMAL | + | + | + | + | − |
| PLANT | + | + | − | − | − |
| SUBSTANCE | + | − | − | − | − |



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

ity on transmission of natural and nonnatural concepts was Bartlett's (1932) study of how British university students remembered and then transmitted a culturally unfamiliar story (a Native American folk tale). Over successive retellings of the story, some culturally unfamiliar items or events were dropped. Other unfamiliar items were distorted, being replaced by more familiar items (e.g., a canoe replaced by a rowboat). Bartlett reasoned that items inconsistent with students' cultural expectations were harder to represent and recall, hence less likely to be transmitted than items consistent with expectations.

Recent studies, however, suggest that under some conditions counterintuitive beliefs are *better* recalled relative to intuitive beliefs (Boyer & Ramble 2001). Barrett and Nyhof (2001) asked people to remember and retell Native American folk tales containing natural as well as nonnatural events or objects. Content analysis showed that participants remembered 92% of minimally counterintuitive items, but only 71% of intuitive items.[14]

Although suggestive, these studies leave several issues unresolved. For one: Why don't minimally counterintuitive concepts occupy most of the narrative structure of religions, folktales, and myths? Even casual perusal of culturally successful materials, like the Bible, Hindu Veda, or Maya Popul Vuh, suggests that counterintuitive concepts and occurrences are a minority. The Bible is a succession of mundane events – walking, eating, sleeping, dreaming, copulating, dying, marrying, fighting, suffering storms and drought, and so on – interspersed with a few counterintuitive occurrences, such as miracles and appearances of supernatural agents like God, angels, and ghosts.

An answer to this puzzle may lie in examining memorability for an entire set of beliefs taken as a single unit of transmission, rather than individual beliefs. Accordingly, we conducted a study to examine the memorability of intuitive (INT) and minimally counterintuitive (MCI) beliefs and belief sets over a period of a week (see Table 2 for examples). One group of 44 U.S. students rated these beliefs on degree of supernaturalness using a 6-point Likert scale. Counterintuitive ideas were viewed as more supernatural than intuitive ones, $t(43) = 14.93$, $p < .001$ ($M = 2.51$ vs. $M = 4.62$). Another group of U.S. students recalled these items over time. INT beliefs showed better recall rates than MCI beliefs, both immediately (Fig. 1) and after a one-week delay (Fig. 2), $F(4,104) = 9.51$, $p < .001$. Because the two kinds of beliefs were matched (each term was equally

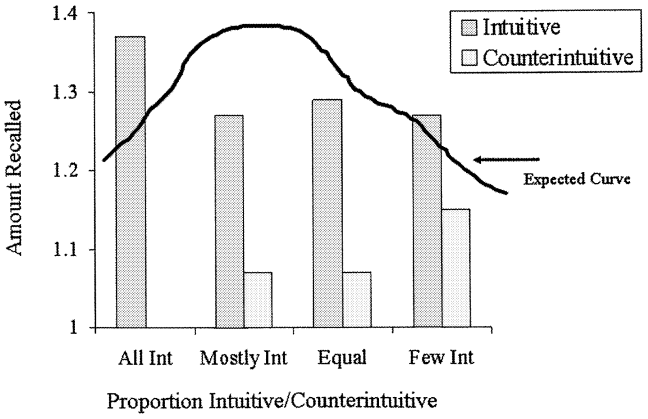



Figure 1. Immediate recall for U.S. students by proportion of intuitive (INT) and minimally counterintuitive (MCI) beliefs.

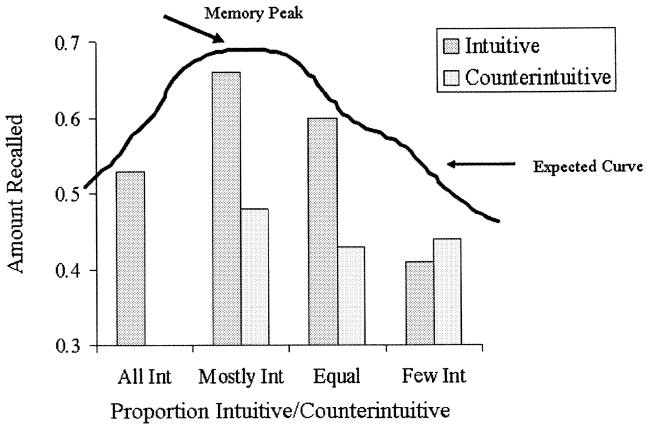



Figure 2. Delayed 1-week recall for U.S. students by proportion of intuitive (INT) and minimally counterintuitive (MCI) beliefs.

likely to occur in an intuitive and counterintuitive belief), the intuitiveness factor, not other unknown factors left to vary, contributed to the recall advantage of the intuitives (Norenzayan & Atran 2004; Norenzayan et al. 2005).

We replicated this finding with a different set of ideas, where a sharper distinction was made between counterintuitive ideas and ideas that are intuitive but bizarre, and between degrees of counterintuitiveness. A group of 107 U.S. participants from another university received ideas that were (1) intuitive and ordinary (INT), (2) intuitive but bizarre (BIZ), (3) minimally counterintuitive (MCI), and (4) maximally counterintuitive (MXCI). Two-word or three-word statements that represented INT, BIZ, MCI, MXCI-Control, and MXCI beliefs were generated (Table 2). Each statement consisted of a concept and one or two properties that modified it. INT statements were created by using a property that was appropriate to the ontological category (e.g., closing door). BIZ statements were created by modifying the concept with an intuitive but bizarre property (e.g., blinking newspaper). MCI statements were created by modifying with a property transferred from *another* ontological category (e.g., thirsty door). Finally, MXCI statements were created by modifying a concept with two properties taken from another ontological category (e.g., squinting wilting brick). For each MXCI statement, a matching statement was generated, only one of the properties being counterintuitive (e.g., chattering climbing pig). Participants received one of two different versions.

Intuitive ideas (INT) had the highest recall; maximally counterintuitive ideas (MXCI), the lowest (Fig. 3). Most distortions occurred within the same ontological category (39 items, or 55%), the majority being within the minimally counterintuitive (MCI) category (23 items = 59% of all same-category distortions). For example, "cursing horse" was remembered as "laughing horse" (both MCI). For distortions that crossed ontological boundaries, the most common was from counterintuitive to intuitive (14 distortions = 20%). The least common distortion was from intuitive to counterintuitive: Only one such distortion was found (1.4%). Results for distorted items, with a preference for rendering counterintuitive beliefs intuitive, follows the main lines of Bartlett's (1932) study.

One finding that converges with previous studies was that minimally counterintuitive beliefs degraded at a lower


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

Table 2. *Examples of intuitive statements (INT) and bizarre (BIZ), and minimally counterintuitive (MCI) and maximally counterintuitive (MXCI) counterparts in counterbalanced design*

| Version 1 | Version 2 |
|---|---|
| INT | INT |
| 1.Crumbling Ice | 1.Crystallizing Ice |
| 2.Crystallizing Glass | 2.Crumbling Glass |
| 3.Gossiping Child | 3.Chanting Child |
| 4.Chanting Man | 4.Gossiping Man |
| 5.Grazing Cow | 5.Wandering Cow |
| 6.Wandering Deer | 6.Grazing Deer |
| BIZ | BIZ |
| 7.Nauseating Cat | 7.Dangling Cat |
| 8.Dangling Squirrel | 8.Nauseating Squirrel |
| 9.Blinking Newspaper | 9.Floating Newspaper |
| 10.Floating Pencil | 10.Blinking Pencil |
| MCI | MCI |
| 11.Giggling seaweed | 11.Sobbing Seaweed |
| 12.Sobbing Oak | 12.Giggling Oak |
| 13.Cursing Horse | 13.Admiring Horse |
| 14.Admiring Frog | 14.Cursing Frog |
| 15.Solidifying Lady | 15.Melting Lady |
| 16.Melting Grandfather | 16.Solidifying Grandfather |
| MXCI-Control | MXCI-Control |
| 17.Cheering Limping Turtle | 17.Chattering Climbing Turtle |
| 18.Chattering Climbing Pig | 18.Cheering Limping Pig |
| MXCI | MXCI |
| 19.Squinting Wilting Brick | 19.Squealing Flowering Brick |
| 20.Squealing Flowering Marble | 20.Squinting Wilting Marble |

rate after immediate recall. Minimally counterintuitive beliefs may have a potent survival advantage over intuitive beliefs: once processed and recalled, they degrade less than intuitive ones. Disadvantage in recall may be offset by resilience, so that over generations of transmission, an idea that is less remembered, but also less degradable, may prevail over an idea that is initially remembered well but eventually dies out because of a higher rate of degradation.

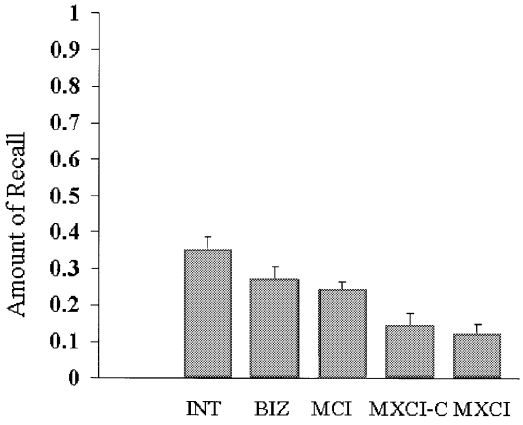



Figure 3. Delayed 1-week recall for U.S. students by proportion of intuitive (INT), bizarre (BIZ), minimally counterintuitive (MCI) and maximally counterintuitive (MXCI) beliefs (error bars, 95% confidence interval).

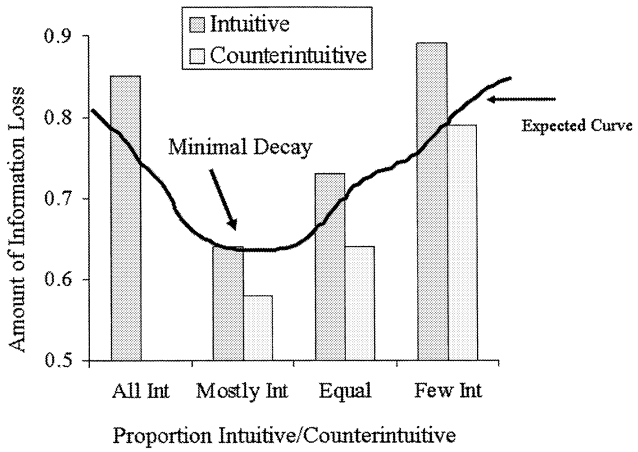



Figure 4. Memory degradation over 1-week (immediate minus delayed recall) for U.S. students by proportion of intuitive (INT) and minimally counterintuitive (MCI) beliefs.

As to belief sets, the one that was mostly intuitive, combined with a few minimally counterintuitive ones, had the highest rate of delayed recall and the lowest rate of memory degradation over time (Fig. 4).[15] This is the recipe for a successful transmission of cultural beliefs, and it is the cognitive template that characterizes most popular folktales and religious narratives. Critically, the belief set with a majority of minimally counterintuitive beliefs had the lowest rate of delayed recall and highest level of memory degradation. In fact, this is a cognitive template rarely encountered in existing culturally successful materials. Thus, the way natural and non-natural beliefs are combined is crucial to the survival of a cultural ensemble of beliefs, such as those that form the core of any religious tradition.

With Yukatek Maya speakers we found the same recall pattern as in the U.S. follow-up. Also, minimally counterintuitive beliefs were again more resilient than intuitive ones, confirming the U.S. pattern. Finally, we found no reliable differences between the Yukatek recall pattern after one week and after three months (Fig. 5). These results indicate cultural stabilization of that pattern.

In sum, minimally counterintuitive beliefs, as long as they come in small proportions, help people remember and

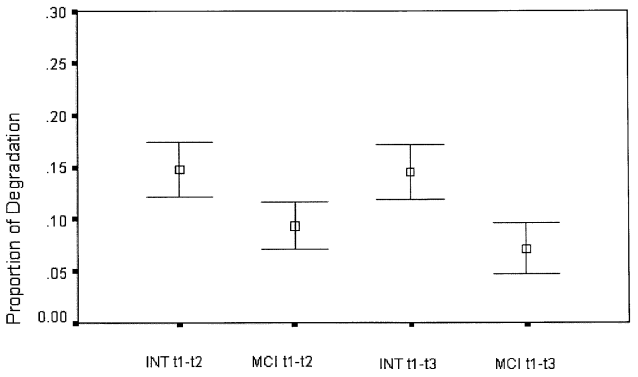



Figure 5. Memory degradation over 1 week (t1–t2) and over 3 months (t1–t3) for Yukatek Maya by proportion of intuitive and minimally counterintuitive beliefs (error bars, 95% confidence interval).



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

presumably transmit the intuitive statements. A small proportion of minimally counterintuitive beliefs gives the story a mnemonic advantage over stories with no counterintuitive beliefs or with far too many counterintuitive beliefs, just as moderately spiced-up dishes have a cultural advantage over bland or far too spicy dishes. This dual aspect of supernatural beliefs and belief sets – commonsensical and counterintuitive – renders them intuitively compelling yet fantastic, eminently recognizable but surprising. Such beliefs grab attention, activate intuition, and mobilize inference in ways that greatly facilitate their mnemonic retention, social transmission, cultural selection, and historical survival.

## 5. Metarepresenting counterintuitive worlds: A Theory of Mind experiment

Thus far we have claimed that the presence of minimally counterintuitive beliefs in religious belief sets favors the production, transmission, and cultural survival of those belief sets over time. We have also provided initial experimental support for the claim, although clearly much more needs to be done. This claim leaves open the issue of how counterintuitive beliefs can be formed at all. If counterintuitive beliefs arise by violating innately given expectations about how the world is built, how can we possibly bypass our own hardwired concepts to form counterintuitive religious beliefs? The answer is that we don't entirely bypass commonsense understanding but conceptually parasitize it to transcend it. This occurs through the species-specific cognitive process of metarepresentation.

Humans have a metarepresentational ability, that is, they form representations of representations. This ability allows people to understand a drawing or picture of someone or something as a drawing or picture and not the real thing. It lets us enjoy novels and movies *as fiction* that can emotionally arouse us without actually threatening us. It lets us think about being in different situations and deciding which are best for the purposes at hand, without our having to actually live through (or die in) the situations we imagine. It affords us the capacity to model the world in different ways, and to conscientiously change the world by entertaining new models that we invent, evaluate, and implement. It enables us to become aware of our experienced past and imagined future as past or future events that are distinct from the present that we represent to ourselves, and so permits us to reflect on our own existence. It allows people to comprehend and interact with one another's minds.

Equally important for our purposes, metarepresentation allows humans to retain half-understood ideas (Atran & Sperber 1991; Sperber 1985). By embedding half-baked (quasi-propositional) ideas in other factual and commonsense beliefs, these ideas can simmer through personal and cultural belief systems and change them. Children come to terms with the world in similar ways when they hear a new word. A half-understood word is initially retained metarepresentationally, as standing in for other ideas the child already has in mind. Initially, the new word is assigned an ontological category: For example, if "andro chases balls," then andro must be an *animal* or *person,* like Fido or Fred.

After Dennett (1978), most researchers in folkpsychology, or "theory of mind" (ToM), maintain that attribution of mental states, such as belief and desire, to other persons requires metarepresentational reasoning about false beliefs. Only when the child can understand that other people's beliefs are *only* representations – not just recordings of the way things are – can the child entertain and assess other people's representations as veridical or fictional, truly informative or deceptive, exact or exaggerated, worth changing one's own mind for or ignoring. Only then can the child appreciate that God thinks differently from most people, in that only God's beliefs are always true.

In one of the few studies to replicate findings on ToM in a small-scale society (cf. Avis & Harris 1991), Knight et al. (2004) showed 48 Yukatek-speaking children (26 boys, 22 girls) a tortilla container and told them, "Usually tortillas are inside this box, but I ate them and put these shorts inside." They asked each child in random order what a person, God, the sun (*k'in*), principal forest spirits (*yumil k'ax'ob',* "Masters of the Forest"), and other minor spirits (*chiichi'*) would think was in the box. As with American children (Barrett et al. 2001), the youngest Yukatek (4 years) overwhelmingly attributed true beliefs to both God and people in equal measure. After age 5, the children attributed mostly false beliefs to people but attributed mostly true beliefs to God (Fig. 6).[16] Thus, 33% of the 4-year-olds said that people would think tortillas were in the container versus 77% of 7-year-olds. In contrast, no significant correlation was detected between answers for God and age, $r(46) = .06$.

Collapsing over ages, Yukatek children attributed true beliefs according to a hierarchy of human and divine minds, one in which humans and minor spirits are seen as easier to deceive. Mental states of humans were perceived as different from those of God ($Z = 3.357$, $p = .001$) and those of Masters of the Forest and the Sun Deity ($Z = 1.89$, $p = .06$ for both). God was seen as all-knowing, and local religious entities were somewhere in between (Fig. 7).

In a follow-up with 7 female and 7 male Itza' Maya adults in Guatemala's Peten rainforest, all (but one man) responded that God had a true belief because, as several respondents stated, "He can see through the basket as if it were transparent." All (save the same man) thought a person coming upon the basket would have a false belief about its contents. Six men and four women thought the forest spirits (*arux*) would know the basket's true contents. Overall (for men as well as women), mental states of humans




Figure 6. Percentage of Yukatek Maya children's attributions of false beliefs to God and persons by age.


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

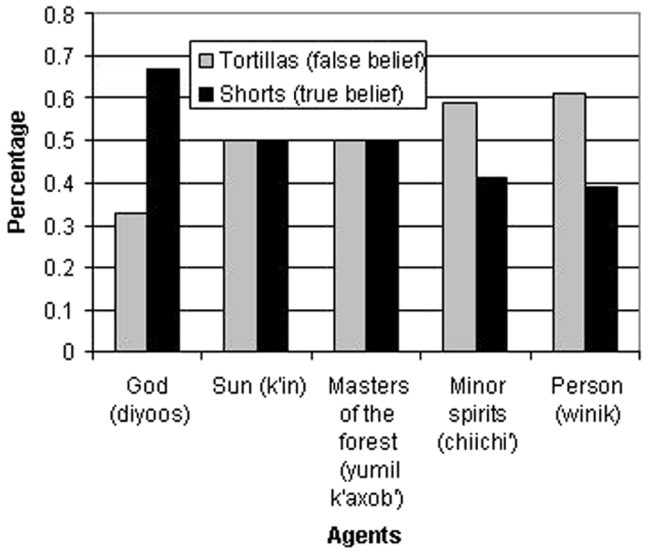



Figure 7. Percentage of responses (all groups) Maya children's attribution of false or true beliefs to persons and various supernatural agents.

were perceived as different from those of God ($Z = 3.207$, $p = .001$) and forest spirits ($Z = 3.000$, $p = .003$), but God and forest spirits were not significantly different from one another. For Itza' adults (but significantly more so for men who venture into the forest than women who generally do not) beliefs in forest spirits have measurable behavioral consequences for biodiversity, forest sustainability, and so forth (Atran et al. 2002). In brief, from an early age people may reliably attribute to supernaturals cognitions that they believe are *different* and *truer* than those attributed to humans.

## 6. From false belief to costly commitment

In this section, we argue that the human metarepresentational ability to deceive and defect has been managed by communicative displays of passionate commitment to omniscient supernatural agents, who unlike humans do not succumb to false beliefs and thus can act as guarantors for future in-group cooperation. Expression of religious prescriptions performatively signals and establishes cognitive and emotional commitment to seek convergence, but it doesn't specify (the propositional content of) *what* people should converge to. The truth about religious prescriptions is accepted on faith and communicated through ritual display, not discovered or described as a set of factual or logical propositions. The result of such convergence is to perpetuate a stable community of cooperators who sacrifice for the group in the short run, but benefit from it in the long run.

One plausible evolutionary story is that understanding agency, together with metarepresenting false belief and deceit, emerged as a later development of intentional communicative displays that signaled possibilities for hominids to cooperate (or deceive) in a wide variety of situations (Leslie & Frith 1987). Autistic children, who selectively fail at false belief tasks, seem to miss intentional communicative display. Although they can often imitate a gesture, and so represent it, they can't go beyond this primary representation to infer that the gesture stands for something else. Thus, unlike non-autistic 1-year-olds (Baron-Cohen 1995; Masur 1983), older autistic children can't signal communicative intent by pointing (as only humans can; see Premack & Woodruff 1978; Povinelli 2000). They can't metarepresent the relation *intentionally communicate,* between a person as an agent (mother), a stimulus situation (upturned palm oriented towards a vase of flowers), and an inferred situation (child giving flowers to mother). Neither, apparently, can they entertain counterfactual beliefs. This can be particularly striking in children suffering from Asperger's Syndrome, a high-functioning form of autism.[17]

Religious acts of faith incorporate universal, metarepresentational features of pragmatic communication, including: *pretend* (that $p$) and *promise* (to do $p$). These are *social* acts common to all normally interacting human agents. A principal difference between religious and nonreligious employments of these behaviors is that the situation that is represented ($p$) in a religious act is not a state of affairs by which the truth, adequacy, or accurateness of the representation is evaluated. Rather, a religious representation (statement or other display) is *always right,* and the situation to which it is properly applied *is made* to conform to what is conventionally stipulated to be the case.

In pretense, a person believes that [$p$ is false] because not-$p$ is demonstrably or verifiably the case. In faith, a person believes that [$p$ is true] because $p$ is the Word of God, and because God always speaks the truth. Faith, like pretense, necessarily involves metarepresentation, namely, the representation in the brackets, where $p$ is metarepresented. In pretense, though, $p$'s content is well understood, and the state of affairs it represents is assessable by observation for truth or falsity (e.g., $p$ = "this banana is a telephone"). In faith, $p$ is not well understood, and the supposed state of affairs it represents cannot be assessable by observation (e.g., $p$ = "this wine is Christ's blood") (Sperber 1975b; cf. Ayer 1950 on religious "pseudo-propositions"). Nonetheless, because the word of god is always true, religious believers are not concerned with whether $p$ is true or not, but with what $p$, which is true, could possibly mean (connote) for them in each situation.

As with pretense, religious acts of faith involve exaggerated gestures that are intended to connote a situation that goes beyond the one perceptually manifest. For example, the act of receiving the host during Mass is an extraordinary eating display, where people are typically fed on their knees with no chewing of the wafer allowed. It is obvious to everyone that the intended goal of the display is not eating, but communion (Rappaport 1999). The meaning of an act of faith like communion is not an inference to specific propositions, but to an emotionally charged network of partial and changeable descriptions of counterfactual and counterintuitive worlds.

In sum, human metarepresentational abilities, which are intimately bound to fully developed cognitions of agency and intention, also allow people to entertain, recognize, and evaluate the differences between true and false beliefs. Given the ever-present menace of enemies within and without, concealment, deception, and the ability to both generate and recognize false beliefs in others would favor survival. But because human representations of agency and intention include representations of false belief and deception, human society is forever under threat of moral defection.

If some better ideology is likely to be available some-



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

where down the line, then, reasoning by backward induction, there is no more justified reason to accept the current ideology than convenience. As it happens, the very same metacognitive aptitude that initiates this problem also provides a resolution through metarepresentation of minimally counterintuitive worlds. Invoking supernatural agents who may have true beliefs that people ordinarily lack creates the arational conditions for people to steadfastly commit to one another in a moral order that goes beyond apparent reason and self-conscious interest. In the limiting case, an omniscient and omnipotent agent (e.g., the supreme deity of the Abrahamic religions) can ultimately detect and punish cheaters, defectors, and free riders no matter how devious (cf. Frank 1988; Dennett 1997).

In the competition for moral allegiance, secular ideologies are at a disadvantage. For, if people learn that all apparent commitment is self-interested convenience or worse, manipulation for the self-interest of others, then their commitment is debased and withers. Especially in times of vulnerability and stress, social deception and defection in the pursuit of self-preservation is therefore more likely to occur, as Ibn Khaldun recognized centuries ago (Ibn Khaldun 1318/1958, II, iii, p. 41). Religion passionately rouses hearts and minds to break out of this viciously rational cycle of self-interest and to adopt group interests that may benefit individuals in the long run. Commitment to the supernatural underpins the "organic solidarity" (Durkheim 1912/1995) that makes social life more than simply a contract among calculating individuals. Commitment to the supernatural is further sustained by the relieving of pervasive existential anxieties, to which we now turn.

## 7. Existential anxiety: A motivation experiment

If supernatural agents are cognitively salient and possess omniscient and omnipotent powers, then they can be invoked to ease existential anxieties such as death and deception that forever threaten human life everywhere. This section summarizes experiments (cf. Norenzayan & Hansen 2005), linking adrenaline-activating death scenes to increased belief in God's existence and the efficacy of super-

Table 3. *Three stories with matching events used to prime feelings of religiosity: Neutral (uneventful), Death (stressful), Religious (prayer scene)*

| | Neutral | Death | Religious |
|---|---|---|---|
| 1 | A mother and her son are leaving home in the morning. | A mother and her son are leaving home in the morning. | A mother and her son are leaving home in the morning. |
| 2 | She is taking him to visit his father's workplace. | She is taking him to visit his father's workplace. | She is taking him to visit his father's workplace. |
| 3 | The father is a laboratory technician at Victory Memorial Hospital. | The father is a laboratory technician at Victory Memorial Hospital. | The father is a laboratory technician at Victory Memorial Hospital. |
| 4 | They check before crossing a busy road. | They check before crossing a busy road. | They check before crossing a busy road. |
| 5 | While walking along, the boy sees some wrecked cars in a junk yard, which he finds interesting. | While crossing the road, the boy is caught in a terrible accident, which critically injures him. | While walking along, the boy sees a well-dressed man stop by a homeless woman, falling on his knees before her, weeping. |
| 6 | At the hospital, the staff are preparing for a practice disaster drill, which the boy will watch. | At the hospital, the staff prepares the emergency room, to which the boy is rushed. | At the hospital, the boy's father shows him around his lab. The boy listens politely, but his thoughts are elsewhere. |
| 7 | An image from a brain scan machine used in the drill attracts the boy's interest. | An image from a brain scan machine used in a trauma situation shows severe bleeding in the boy's brain. | An image from a brain scan that he sees reminds him of something in the homeless woman's face. |
| 8 | All morning long, a surgical team practices the disaster drill procedures. | All morning long, a surgical team struggles to save the boy's life. | On his way around the hospital, the boy glances into the hospital's chapel, where he sees the well-dressed man sitting alone. |
| 9 | Make-up artists are able to create realistic-looking injuries on actors for the drill. | Specialized surgeons are able to reattach the boy's severed feet, but cannot stop his internal hemorrhaging. | With elbows on his knees, and his head in his hands, the man moves his lips silently. The boy wants to sit beside him, but his father leads him away. |
| 10 | After the drill, while the father watches the boy, the mother leaves to phone her other child's preschool. | After the surgery, while the father stays by the dead boy, the mother leaves to phone her other child's preschool. | After a brief tour of the hospital, while the father watches the boy, the mother leaves to phone her other child's preschool. |
| 11 | Running a little late, she phones the preschool to tell them she will soon pick up her child. | Barely able to talk, she phones the preschool to tell them she will soon pick up her child. | Running a little late, she phones the preschool to tell them she will soon pick up her child. |
| 12 | Heading to pick up her child, she hails a taxi at the number nine bus stop. | Heading to pick up her child, she hails a taxi at the number nine bus stop. | Heading to pick up her child, she hails a taxi at the number nine bus stop. |


https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

natural intervention in human affairs. The experiment is also aimed at commitment theories of religion that neglect special attention to the supernatural.

Our experiment builds on a study by Cahill and colleagues dealing with the effects of adrenaline (adrenergic activation) on memory (Cahill et al. 1994). They showed college students a series of slides and a storyline about a boy riding a bike. Some subjects were exposed to an uneventful story: The boy rides his bike home, and he and his mother drive to the hospital to pick up his father (who is a doctor). For the other participants, the story begins and ends in much the same way, but the middle is very different: The boy is hit by a car and rushed to the hospital's emergency room, where a brain scan shows severe bleeding from the boy's brain and specialized surgeons struggle to reattach the boy's severed feet. After exposure to the stories, and before being tested for recall, half the subjects were given either a placebo pill or a drug (propranolol) that blocks the effects of adrenaline. The placebo and drug groups recalled the uneventful story equally well. Only the placebo group, however, remembered the emotional story more accurately than the uneventful one.

Our hypothesis was that existential anxieties (particularly about death) not only deeply affect how people remember events but also their propensity to interpret events in terms of supernatural agency. We primed each of three groups of college students with one of three different stories (Table 3): Cahill et al.'s uneventful story (neutral prime), Cahill et al.'s stressful story (death prime), and another uneventful story whose event-structure matched the other two stories but which included a prayer scene (religious prime). Afterwards, each group of subjects read a *New York Times* article (2nd October 2001) whose lead ran: "Researchers at Columbia University, expressing surprise at their own findings, are reporting that women at an in vitro fertilization clinic in Korea had a higher pregnancy rate when, unknown to the patients, total strangers were asked to pray for their success." The article was given under the guise of a story about "media portrayals of scientific studies." Finally, students rated strength of their belief in God and the power of supernatural intervention (prayer) on a 9-point scale.

Results show that strength of belief in God's existence (Fig. 8), and in the efficacy of supernatural intervention (Fig. 9), are reliably stronger after exposure to the death prime than after the neutral and religious primes, $F(1, 74) = 7.44$, $p < .01$, and $F(1, 74) = 3.88$, $p = .05$, respectively (no significant differences between either uneventful story). This effect held even after controlling for religious background and prior degree of religious identification.

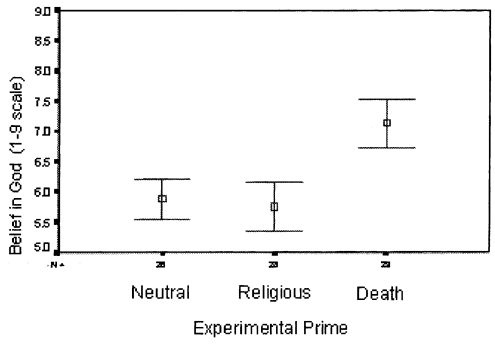



Figure 8. Strength of belief in God's existence after priming (neutral, religious, or death) (error bars, 95% confidence interval).

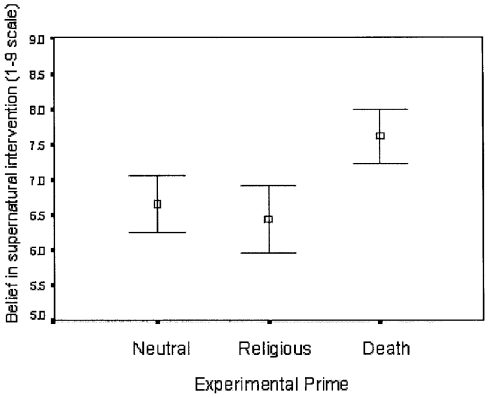



Figure 9. Strength of belief in supernatural power of prayer after priming (neutral, religious, or death) (error bars, 95% confidence interval).

Terror Management Theory (TMT) maintains that cultural worldview is a principal buffer against the terror of death. Accordingly, TMT experiments show that thoughts of death induce people to reinforce their cultural (including religious) worldview and derogate alien worldviews (Greenberg et al. 1990; Pyszczynski et al. 1999). According to the worldview-defense hypothesis, then, awareness of death should enhance belief in a culturally familiar deity, but diminish belief in a culturally unfamiliar deity. Our view suggests that the need for belief in supernatural agency is possibly a qualitatively distinct buffer against the terror of death that overrides worldview defense needs.

To test this idea, in a follow-up, 73 American undergraduates were told that the prayer groups were Buddhists in Taiwan, Korea, and Japan. Supernatural belief was measured either shortly after the primes, or after a significant delay between the primes and the belief measures. When the primes were recently activated, as expected there was a stronger belief in the power of Buddhist prayer in the death prime than in the control prime, $F(1, 33) = 6.65$, $p = .01$. Remarkably, the mostly Christian death-primed subjects who previously self-identified as strong believers in their religion were *more* likely to believe in the power of Buddhist prayer, $r(16) = .68$, $p < .01$. In the neutral (control) condition, there was no correlation between religious identification and belief in Buddhist prayer among a similar group of mostly Christian subjects. Given a choice between supernatural belief versus rejecting an alien worldview (Buddhism), Christians chose the former. This finding is difficult to explain in terms of cultural worldview bolstering, but it possibly reflects a strong belief in immortality as a buffer against death, as articulated in TMT (Solomon et al. 1991).

In a cross-cultural extension, 75 Yukatek-speaking Maya villagers were tested, using stories matched for event structure but modified to fit Maya cultural circumstances. They were also asked to recall the priming events. We found no differences among primes for belief in the existence of God and spirits (near ceiling in this very religious society). However, subjects' belief in the efficacy of prayer for invoking the deities was significantly greater with the death prime than with religious or neutral primes, $\chi^2(2, N = 75) = 10.68$, $p = .005$. Awareness of death more strongly motivates religiosity than mere exposure to emotionally nonstressful religious scenes, like praying. This supports the



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

claim that emotionally eruptive existential anxieties motivate supernatural beliefs.

We found no evidence for differences in recall of priming events after subjects rated their strength of belief in God and the efficacy of supernatural intervention. With this in mind, note that uncontrollable arousal mediated by adrenergic activation (e.g., subjects chronically exposed to death scenes) can lead to posttraumatic stress syndrome if there is no lessening of terror and arousal within hours; however, adrenergic blockers (e.g., propranolol, guanfacine, possibly antidepressants) can interrupt neuronal imprinting for long-term symptoms, as can cognitive-behavioral therapy (work by Charles Marmar discussed in McReady 1999, p. 9). Heightened expression of religiosity following exposure to death scenes that provoke existential anxieties may also serve this blocking function (Atran 2002b). We plan to test the further claim that existential anxieties not only spur supernatural belief, but that these beliefs are in turn affectively validated by assuaging the very emotions that motivate belief in the supernatural.

All of this isn't to say that *the* function of religion is to promise resolution of all outstanding existential anxieties any more than *the* function of religion is to neutralize moral relativity and establish social order, to give meaning to an otherwise arbitrary existence, to explain the unobservable origins of things, and so forth. Religion has no evolutionary function per se. It is rather that existential anxieties and moral sentiments constitute – by virtue of evolution – ineluctable elements of the human condition; and that the cognitive invention, cultural selection, and historical survival of religious beliefs have resulted, in part, from success in accommodating these elements. There are other factors in this success, involving naturally selected elements of human cognition, such as the inherent susceptibility of religious beliefs to modularized (innate and universal) conceptual and mnemonic processing.

## 8. Conclusion: Evolution's canalizing landscape

Think metaphorically of humankind's evolutionary history as a landscape formed by different mountain ridges. This landscape functions everywhere to canalize, but not determine, individual and cultural development. It greatly reduces the possible sources of religious expression into structures that constantly reappear across history and societies.

This landscape is shaped by natural selection. It is ancestrally defined by specific sets of affective, social, and cognitive features – different mountain ridges. Each ridge has a distinct contour, with various peaks whose heights reflect evolutionary time. One such evolutionary ridge encompasses panhuman emotional faculties, or "affect programs." Some of these affect programs, such as surprise and fear, date at least to the emergence of reptiles. Others, such as grief and guilt, may be unique to humans. Another ridge includes social-interaction schema. Some schema may go far back in evolutionary time, such as those involved in detecting predators and seeking protectors, or which govern direct "tit-for-tat" reciprocity ("you scratch my back, I'll scratch yours"). Other social-interaction schema seem unique to humans, such as committing to non-kin. Still another ridge encompasses panhuman mental faculties, or cognitive *modules*, like folkmechanics, folkbiology, folkpsychology. Folkmechanics is this ridge's oldest part, with links to amphibian brains. Folkpsychology is the newest, foreshadowed among apes. Only humans appear to metarepresent multiple models of other minds and worlds (Tomasello et al. 1993), including the supernatural.

Human experience lies along this evolutionary landscape, usually converging on more or less the same life paths – much as rain that falls anywhere in a mountain-valley landscape, drains into a limited set of lakes or rivers (Kauffman 1993; Sperber 1996). As humans randomly interact and "walk" through this landscape, they naturally tend towards certain forms of cultural life, including religious paths. Cultures and religions don't exist apart from the individual minds that constitute them and the environments that constrain them, any more than a physical path exists apart from the organisms that tread and groove it and the surrounding ecology that restricts its location and course. Individual minds mutually interact within this converging landscape in an open-ended time horizon, exploiting its features in distinctive ways. The result is socially transmitted amalgamations that distinctively link landscape features with cognitive, affective, and interactional propensities. This produces the religious and cultural diversity we see in the world and throughout human history.

Nevertheless, all religions follow the same structural contours. They invoke supernatural agents to deal with emotionally eruptive existential anxieties, such as loneliness, calamity, and death. They have malevolent and predatory deities as well as more benevolent and protective ones. These systematically, but minimally, violate modularized expectations about folkmechanics, folkbiology, and folkpsychology. And religions communally validate counterintuitive beliefs through musical rituals and other rhythmic coordinations of affective body states. Finally, these landscape features are mutually constraining. They include evolved constraints on emotional feelings and displays, modularized conceptual and mnemonic processing, and social commitments that attend to information about cooperators, protectors, predators, and prey.

NOTES

**1.** We make no conceptual distinction between "culture" and "society" or "mind" and "brain."

**2.** This framework is also informed by the first author's (Atran's) ethnographic sojourns among Lowland Maya (Mesoamerica), Druze mountaineers (Middle East), Pashtun nomads (Central Asia), Tamil Hindu farmers (South India), and Ladakhi Buddhist tanshumants (Himalaya), and by the second author's (Norenzayan's) familiarity with the religious civil wars of Lebanon (1975–1991).

**3.** Evolutionarily, at least some basic emotions preceded conceptual reasoning: surprise, fear, anger, disgust, joy, sadness (Darwin 1872/1965; Ekman 1992). These may have further evolved to incite reason to make inferences about situations relevant to survival decisions. Existential anxieties are by-products of evolved emotions, such as fear and the will to stay alive, and of evolved cognitive capacities, such as episodic memory and the ability to track the self and others over time. For example, because humans are able to metarepresent their own selves and mentally travel in time (Wheeler et al. 1997), they cannot avoid overwhelming inductive evidence predicting their own death and that of persons to whom they are emotionally tied, such as relatives, friends, and leaders. Emotions compel such inductions and make them salient and terrifying. This is "The Tragedy of Cognition." All religions propose a supernatural resolution in some minimally counterfactual afterlife.

**4.** Although the Buddha and the buddhas are not regarded as gods, Buddhists clearly conceive of them as "counter-intuitive agents" (Pyysiännen 2003). The Chinese Buddhist Pantheon in-



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

cludes the 18 *Lohan,* or supernatural guardian angels known for their great wisdom, courage, and supernatural power, and the four *Si-Ta-Tien-Wang,* or Guardian Kings of the four directions (akin to the Maya *Chaak*). In Sri Lanka, the Sinhalese relics of the Buddha have miraculous powers. In India, China, Japan, Thailand, and Vietnam, there are magic mountains and forests associated with the Buddha; and the literature and folklore of every Buddhist tradition recount amazing events surrounding the Buddha and the buddhas.

**5.** Experiments with adults in the United States (Barrett & Keil 1996) and India (Barrett 1998) further illustrate the gap between theological doctrine and actual psychological processing of religious concepts. When asked to describe their deities, subjects in both cultures produced abstract and consensual theological descriptions of gods as being able to do anything and anticipate and react to everything at once; always knowing the right thing to do; and being able to dispense entirely with perceptual information and calculation. When asked to respond to narratives about these same gods, the same subjects described the deities as being in only one place at a time, puzzling over alternative courses of action, and looking for evidence in order to decide what to do (e.g., to first save Johnny, who's praying for help because his foot is stuck in a river in the United States, and the water is rapidly rising; or to first save little Mary, whom He has seen fall on railroad tracks in Australia where a train is fast approaching).

**6.** One distinction between fantasy and religion is knowledge of its source. People know or assume that public fictions (novels, movies, cartoons, etc.) were created by specific people who had particular intentions for doing so. Religious believers assume that utterances or texts connected with religious doctrines are authorless, timeless, and true. Consequently, they don't apply ordinary criteria of relevance to religious communications to figure out the speaker's true intentions or check on whether God is lying or lacking information (Sperber & Wilson 1986).

**7.** As Dan Sperber (1996) asked in an open communication to the Evolution and Human Behavior Society: "Is fitness a matter of having descendants with a recognizable ideology? Of population size? Of variations in size (expansion)? Of duration? Of some weighted combination of size and duration? What of social systems that expand rapidly at the expense of heritability (empires)?" Without answers to such questions (and none seem forthcoming) the idea of societal-level fitness is hopelessly vague.

**8.** For each *natural domain,* there is a proper domain and (possibly empty) actual domain (Sperber 1994). A *proper domain* is information that is the cognitive module's naturally selected function to process. The *actual domain* of a module is any information in the organism's environment that satisfies the module's input conditions whether or not the information is functionally relevant to ancestral task demands – that is, whether or not it also belongs to its proper domain. For example, cloud formations and unexpected noises from inanimate sources (e.g., a sudden, howling gush of wind) readily trigger inferences to agency among people everywhere. Although clouds and wind occurred in ancestral environments, they had no functional role in recurrent task problems with animate beings. Similarly, moving dots on a screen do not belong to *agency*'s proper domain because they could not have been involved with ancestral task demands. Like clouds and wind, moving dots on computer screens belong to its actual domain. A parallel example is food-catching behavior in frogs. When a flying insect moves across the frog's field of vision, bug-detector cells are activated in the frog's brain. Once activated, these cells in turn massively fire others in a chain reaction that usually results in the frog shooting out its tongue to catch the insect. The bug-detector is primed to respond to any small dark object that suddenly enters the visual field (Lettvin et al. 1961). If flying insects belong to the proper domain of frog's food-catching module, then small wads of black paper dangling on a string belong to the actual domain.

**9.** Psychoanalytic (Erikson 1963; Freud 1913/1990) and attachment (Bowlby 1969; Kirkpatrick 1998) theories hold that primary deities are surrogate parents who assuage existential anxieties. But ethnographic reports indicate that malevolent and predatory deities are as culturally widespread, historically ancient, and socially supreme as benevolent deities. Examples include cannibalistic spirits of small-scale Amazonian, sub-Saharan African, and Australian aboriginal societies, as well as bloodthirsty deities of larger-scale civilizations that practiced human sacrifice, such as Moloch of the Ancient Middle East, the death goddess Kali of tribal Hindus, and the Maya thunder god Chaak. Psychological findings on false-belief tasks (see below) further indicate that beliefs about people are not the basis of beliefs about God because the developmental trajectories of these two belief sets diverge from the outset.

**10.** Another example from ethology offers a parallel. Many bird species have nests parasitized by other species. Thus, the cuckoo deposits eggs in passerine nests, tricking the foster parents into incubating and feeding the cuckoo's young. Nestling European cuckoos often dwarf their host parents (Hamilton & Orians 1965): "The young cuckoo, with its huge gape and loud begging call, has evidently evolved in exaggerated form the stimuli which elicit the feeding response of parent passerine birds. . . . This, like lipstick in the courtship of mankind, demonstrates successful exploitation by means of a 'super-stimulus'" (Lack 1968). Late nestling cuckoos have evolved perceptible signals to *manipulate* the passerine nervous system by initiating and then arresting or interrupting normal processing. In this way, cuckoos are able to subvert and co-opt the passerine's modularized survival mechanisms.

**11.** Aristotle (1963) was the first to point out in his *Categories* that such counterintuitive expressions cannot even be judged false because no set of truth conditions could ever be definitely associated with them. He gave the example of "two-footed knowledge." According to him, "two-footed" could be sensibly (truly or falsely) applied to all animals but not to any sort of knowledge. This is because knowledge falls under the ontological category of *nonsubstantial things,* whereas being two-footed falls under the altogether distinct ontological category of *substantial things.* Trying to put together things from different ontological categories produces a "category mistake." For Aristotle, the world that could be properly described in ordinary Greek was the world that is (nomologically). This led him to conflate the world's ontological structure (what philosophy and science consider to be the ultimate "stuff" composing the world) with the semantic structure of language (the constraints that govern the ordinary relations between words and thoughts). Subsequent philosophers have reinterpreted the notion of a category mistake as a logical or semantic "type confusion" (Pap 1963; Sommers 1963). Cognitive and developmental psychologists have experimentally shown that children across cultures do not violate such categorical constraints on language learning when attempting to learn the meaning of words (Keil 1979; Walker 1992).

**12.** Science, like religion, uses metarepresentation in cosmology building: for example, in analogies where a familiar domain (e.g., solar systems, computers, genetic transmission) is used to model some initially less familiar system (e.g., atoms, mind/brains, ideational transmission). In fact, science and religion may use the same analogies; however, there is a difference in these uses. Science aims to reduce the analogy to factual description, where the terms of the analogy are finally specified, with no loose ends remaining and nothing left in the dark: Atoms are scientifically like solar systems if and only both can be ultimately derived from the same set of natural laws. Whereas science seeks to kill the metaphor, religion strives to keep it poetic and endlessly open to further evocation. In religion, these ideas are never fully assimilated with factual and commonsensical beliefs, like a metaphor that metarepresents the earth as a mother but not quite, or an angel as a winged youth but not quite.

**13.** According to Boyer (1994; 1997; 2000), bodiless supernaturals are counterintuitive because they think and act but lack physical substance. The matter is not so simple. First, experiments with infants and adults indicate that ordinary intuitions about causal agents do not require knowledge or perception of material



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press

substance, only the expectation (perhaps never actually realized) that there *ultimately is* a physical source of intentional action (Csibra et al. 1999). Ontological violations block such expectations being realized *even in principle* (e.g., invisible agents versus heard but unseen beings). They countermand rules for eventual processing, not actual perception. Second, not all mental states are equally bound to ordinary intuitions about bodies. Recent studies indicate that children from 5 years on up more readily attribute epistemic mental states (see, think, know) to beings in the afterlife than psychobiological mental states (hunger, thirst, sleepiness) (Bering & Bjorklund 2002). Ordinary distinctions between mind and body (e.g., dreaming) thus seem to provide at least *some* intuitive support for extraordinary beings with disembodied minds (Hobbes 1651/1901).

**14.** Barrett and Nyhof (2001, p. 79) list as common items: "a being that can see or hear things that are not too far away"; "a species that will die if it doesn't get enough nourishment or if it is severely damaged"; "an object that is easy to see under normal lighting conditions." Such items fall so far below ordinary expectations that communication should carry some new or salient information that Barrett and Nyhof (2001, pp. 82–83) report: "common items were remembered so poorly relative to other items. . . . In some instances of retelling these items, participants tried to make the common property sound exciting or unusual." In other words, some subjects tried to meet minimum conditions of relevance (Sperber & Wilson 1986). For the most part, common items failed these minimum standards for successful communication.

**15.** Highest degradation was observed in the mostly MCI and all INT conditions, conforming to an inverse quadratic function, $F(3, 89) = 4.49$, $p < .05$. Memory degraded least in the Mostly INT condition, and increased as the proportion of MCI beliefs increased, resulting in a linear trend, $F(2, 65) = 3.53$, $p = .06$.

**16.** Only additional evidence could show whether children "continue" to think of God in the same way after they become aware of false beliefs (as Barrett et al. 2001 intimate), or (as seems more likely) come to have different reasons for thinking that God would not be deceived.

**17.** To deal with deficits in counterfactual thinking, St. Paul's Church in Alabama (Trenton Diocese) has a special program for autistics: "The church requires that children who receive Holy Communion be able to recognize the difference between ordinary bread and the Eucharist. . . . The St. Paul's program was designed to teach the difference" (Rev. Sam Sirianni, cited in Raboteau 2000).

# Open Peer Commentary

## Gods are more flexible than resolutions

George Ainslie

*Veterans Affairs Medical Center, Coatesville, PA 19320.*
**George.Ainslie@med.va.gov** **http://www.Picoeconomics.com**

**Abstract:** The target article proposes that "counterintuitive beliefs in supernatural agents" are shaped by cognitive factors and survive because they foster empathic concern and counteract existential dread. I argue that they are shaped by motivational forces similar to those that shape our beliefs about other people; that empathic concern is rewarded in a more elementary fashion; and that a major function of these supernatural beliefs may be to provide a more flexible alternative to autonomous willpower in controlling not only dread but also many other unwelcome urges.

The useful hypotheses in this article include: (1) that religion is a form of motivated belief, that is, that religious beliefs and their attendant practices survive insofar as they serve a purpose; (2) that a principal purpose of religion is to deter "social deception and defection in the pursuit of self-preservation"; (3) that another principal purpose of religion is to control "emotionally eruptive existential anxieties" (sect. 1, para. 7); and (4) that human experience, and religious experience in particular, converges "on more or less the same life paths – much as rain that falls anywhere in a mountain-valley landscape, drains into a limited set of lakes or rivers" (sect. 8, para. 3). The authors present a case for how humans may be innately prepared to construct the supernatural beings that populate most religions, because of people's "hair-triggered" attribution of agency to ambiguous percepts, the increased memorability of "minimally counterintuitive" ideas, and people's ability to imagine counterfactual omniscient personae. However, this article presents little about what incentive people have to construct these beings – only some unsurprising data that subjects value religious ideas more in fear-provoking situations.

I agree that supernatural religion is probably an extension of "emotional mechanisms that evolved for mundane adaptive tasks" (sect. 1, para. 2), and that part of its usefulness is sometimes to control selfishness and emotional eruptions. However, I do not think the authors have specified adequate motivational mechanisms to account for these effects. Part of this problem comes from the inadequacy of how behavioral science has come to imagine self-interest and altruism. Rational self-interest is identified with beating out competitors for resources, and rational altruism merely with taking the long view of this competition so as to identify situations where cooperation will be more profitable, hedonically or genetically, than competition (Dawkins 1989; Frank et al. 1993). Given the human openness to seduction by short-term prospects, altruism is sometimes suggested to require self-control (Rachlin 2002), but the point is still to maximize your own survival resources. The authors are right to reject this "'mind-blind' functionalism" (sect. 1.5); but the role they give to religious belief remains one of controlling an innate tendency toward selfishness, through belief in vigilant gods.

An adequate theory of altruism needs to explain why people start out as highly empathic children (Harris 1987; Zahn-Waxler et al. 1992), who then learn to a variable extent to control empathy as an impulse. That is, why is there a basic self-interest in cultivating vicarious emotional experience, which is then partially displaced by the more "objective" self-interest of (say) economic man? This area is largely terra incognita. Motivational theory has not examined even nonvicarious emotions as rewards until recently (Lewis & Haviland-Jones 2000); they are awkward targets for controlled research, and it is hard even to theorize about rewards that require no specific stimulus and have many of the characteristics of behaviors. However, mounting evidence that all reward-responsive organisms discount delayed rewards proportionally to this delay (hyperbolically) rather than at fixed rates (exponentially; Kirby 1997) suggests one mechanism for vicarious emotional reward, based on the innate impulsiveness that such discounting predicts (Ainslie 1995; 2001, pp. 161–86). I can only summarize it here: Emotions are reward-dependent behaviors that have their own appetites and lead to their own innate rewards, rather than being elicited reflexes. Because of a hyperbolic impatience for their rewards, these behaviors are limited by premature satiation, which causes extinction of deliberately emitted emotions; to stay fresh they must be occasioned by uncontrollable events. Such a contingency makes external occasions for emotion valuable, and these occasions seem especially well paced by the apparent experience of other people. Thus, vicarious reward creates an incentive to help the people whose experiences you choose as occasions for emotion, and to resist temptations to exploit them. The recent discovery of "mirror neurons" that initiate copies of other people's behaviors (Iacoboni et al. 1999) suggests a reason why vicarious experience may stand out from other available occasions for emotion. Whatever the mechanism, empathic engagement with its sometime result of altruism is apparently a primary motivated process.



https://doi.org/10.1017/S0140525X04000172 Published online by Cambridge University Press