Washington University
ERes Cover Sheet

Article Title: <u>Functional Origins of Religious Concepts: Ontological and Strategic</u>
<u>Selection in Evolved Minds</u>

Author: <u>Boyer, P.</u>

Source Title: <u>Journal of the Royal Anthropological Institute</u>

Vol.: <u>6</u>   Issue:___   Date: <u>2000</u>   Pages:<u> 195-214</u>

FUNCTIONAL ORIGINS OF RELIGIOUS CONCEPTS: ONTOLOGICAL AND STRATEGIC SELECTION IN EVOLVED MINDS*

PASCAL BOYER

Centre National de la Recherche Scientifique, Lyons

Culturally successful religious concepts are the outcome of selective processes that make
some concepts more likely than others to be easily acquired, stored and transmitted. Among
the constructs of human imagination, some connect to intuitive ontological principles in
such a way that they constitute a small catalogue of culturally successful supernatural con-
cepts. Experimental and anthropological evidence confirm the salience and transmission
potential of this catalogue. Among these supernatural concepts, cognitive capacities for social
interaction introduce a further selection. As a result, some concepts of supernatural agents
are connected to morality, group identity, ritual and emotion. These typical 'religious' super-
natural agents are tacitly presumed to have access to information that is crucial to social
interaction, an assumption that boosts their spread in human groups.

What is the origin of religious concepts? How come we can find concepts
of supernatural agency more or less the world over, with important recurrent
features? This lecture is a 'progress report', an account of how these previously
intractable questions are now a matter of empirical, indeed experimental
enquiry. What brought about this remarkable change is substantial progress in
our understanding of how human minds work. This allows a naturalistic
account of cultural representations[1] that describes how evolved conceptual
dispositions make humans likely to acquire certain concepts more easily than
others.

Cultural transmission, like other forms of human communication, does not
consist in 'downloading' concepts from one mind to another. It requires infer-
ential processes, whereby people attend to cues in other people's behaviour,
infer their communicative intentions and build concepts on the basis of what
they inferred (Sperber 1996; Tomasello, Kruger & Ratner 1993). As a result,
people constantly create variants of other people's representations. To call some
representations 'cultural' is to point to a relative similarity between repre-
sentations held by members of a particular group. The similarity suggests that
some concepts are *selected* in the transmission process, against a whole variety
of variants that are forgotten, discarded and modified.

Obviously, an indefinite variety of factors contribute to the stability and
recurrence of any particular set of representations in a historical context.

However, the aggregation of many individual acquisition processes 'washes out' such local factors. In the long run and in the comparison of many different human groups, local factors cancel each other out. What we find as recurrent features, over time and between groups, are concepts that, all else being equal, tend to resist distortion better than others. The aggregated result of individual acquisition and communication episodes channels cultures along particular paths, with the result that some concepts are both relatively stable within a group and recurrent among different groups.

In the domain of religious concepts, two different selection processes reduce the domain of culturally fit concepts.

First, not all products of human imagination are equally fit for widespread transmission. Although there are no obvious limits to human imagination, we find that in most human groups supernatural notions (including religious concepts but also folklore, superstition, fiction and fantasy) tend to centre around a small catalogue with recurrent features. This is because human minds are equipped with an *intuitive ontology*, a set of expectations about the kinds of things to be found in the world. Among the indefinitely many concepts individuals can imagine and combine, some connect with this ontology in a particular way. As a result, they stand better chances than other concepts of spreading in a relatively preserved form.

Second, not all culturally fit supernatural concepts are of equal social importance. Some of them are connected to representations of group identity, ritual, morality and social interaction, corresponding to what we usually call 'religion', while others are not. This difference stems from another set of cognitive capacities. Humans are greatly dependent upon cooperation and information about potential cooperators, which creates specific cognitive problems. Humans have a set of *strategic capacities* that handle these problems. Now some supernatural concepts are represented in such a way that they activate strategic capacities. As a result, they are more likely than others to generate high commitment and other psychological and social effects typical of 'religion'.

Religious concepts and intuitive ontology[2]

The present account is about mentally represented concepts that are spread in human groups. These concepts do not always match official descriptions of what people are supposed to think, in this case commonly sanctioned theologies. Perhaps more important, people's concepts are not entirely accessible to conscious inspection. This is why we need and use experimental protocols to uncover their contents and organization.

Three features are generally present in the mental representation of such concepts.

First, supernatural concepts *activate a set of ontological categories* that we know are present in normal minds from an early stage of cognitive development. Objects in the environment are identified as belonging to kind-concepts ('telephone', 'giraffe') but also to ontological categories (PERSON, ARTEFACT, ANIMAL, etc.). We now have much better evidence concerning intuitive ontological categories and associated theories, coming from

developmental psychology, from experimental studies with adults, from neuroscience and from the study of cognitive pathologies.[3] Concepts of imaginary objects and beings are intuitively associated with particular ontological categories too. The concept of 'spirit' activates the category PERSON. If you pray to a particular statue of the Virgin, you are standing or kneeling in front of an ARTEFACT. If you think that some antelopes can disappear at will, you must activate your ANIMAL category to represent these special beings.

Second, supernatural concepts invariably *specify information that violates intuitive expectations* associated with the relevant ontological category. Ontological categories are associated with 'intuitive theories'. For instance, the motion and interaction of solid physical objects is the object of an 'intuitive physics' that develops from the first months after birth. Principles of intuitive physics stipulate that objects have continuous trajectories in space and time, that they collide but do not merge on contact, that unsupported objects fall and so forth. The fact that some objects are identified as members of the categories ANIMAL or PERSON triggers 'intuitive biology' principles. These stipulate that live beings have 'essential' qualities inside them that make them members of one particular species, that they are propelled by internal force, that they belong to mutually exclusive classes in a taxonomic hierarchy. The most important domain of intuitive theoretical principles is 'intuitive theory of mind', a domain of expectations and principles that allow us to represent the behaviour of animals and people as guided by internal representations: beliefs, intentions, emotions. Supernatural concepts describe members of ontological categories that violate intuitive expectations for these categories. Spirits are persons that are mostly invisible and go through physical obstacles. Some statues and other artefacts are special because you can talk to them and they will listen and understand. Some mountains are special because they eat food and digest it (see more illustrations in Boyer 1994). These violations are generally explicit and are transmitted by cultural input.

Third, a supernatural concept also *activates the intuitive expectations that are not violated*, among those associated with the relevant ontological category. By contrast with the features above, this remains generally tacit and need not be acquired via social transmission. For instance, people tacitly represent spirits as having minds. That is, spirits are assumed to perceive events, to remember what they perceived, to have beliefs and form intentions on the basis of their beliefs, and so on. Intuitive psychology is spontaneously extended to spirits because they are identified as a special kind of PERSON. Note that our intuitive theory of mind works very well without us ever representing what its principles are, how it computes intentions from behaviours, and so on.

Mentally represented supernatural concepts are extremely diverse. They vary not just from group to group but also from a member of a group to another and even within the same individual, depending on the cognitive task at hand. But diversity and similarity, obviously, are a matter of explanatory viewpoint. If we focus, not just on the concepts themselves but also on the cognitive processes that allow people to acquire, represent and communicate them, we can see that a great variety of supernatural *concepts* are informed by a small number of *templates*. Templates are just principled ways of handling information from intuitive ontology, with the following structure:

(0) lexical label;
(1) pointer to an ontological category;
(2) violation of expectations, either:
    (a) breach of expectations for that particular category, or
    (b) transfer of expectations from another category;
(3) activation of non-violated expectations for the category;
(4) additional encyclopaedic details, that vary from place to place.

Many different religious concepts may correspond to one template. For instance, this is the template for the concepts of 'spirit' that we find in so many cultures:

(1) an ontological category: PERSON;
(2) a violation of intuitive physics, e.g. spirits are invisible;
(3) activation of non-violated expectations: being persons, spirits, have a mind, they can perceive events, form beliefs, have intentions, etc.;
(4) place-holder for additional (local) detail.

Templates are not 'archetypes' or 'ideal types' or a 'family resemblance' for religious concepts. In fact templates are not concepts at all. They are just *pro-cedures* for the use of information provided by intuitive ontology. The cultural success of concepts seems to depend, not on their specific features, but on what intuitive ontological information is used and how; in other words, which template is used.

This account predicts that there are not that many different templates, simply because there are not that many ontological categories and sets of associated intuitive principles. In other words, there are not that many ways of 'tweaking' intuitive ontology so as to produce supernatural concepts, so that a general 'catalogue of the supernatural' should be rather short. If one sticks with the description of intuitive ontology given above (which may of course change depending on better psychological evidence for categories and intuitive principles), five categories are involved: PERSON, ANIMAL, PLANT, ARTEFACT, NATURAL OBJECT (i.e. non-man-made, non-living parts of the environment like rivers, rocks and mountains) and three main domains of inference: intuitive physical expectations, intuitive biology and intuitive 'theory of mind'. Violations are produced either by *breach* or by *transfer*. A breach contradicts intuitive expectations associated with the ontological category (e.g. a table that suddenly disappears, thereby violating intuitive physics activated by the ARTEFACT category). A transfer extends to a category information that is intuitively associated with another category (e.g. a table that breathes, using biological information associated with the ANIMAL category). Categories, types of inferences and possible violations produce only a small list of templates:

(1) Person + breach of physical expectations;
(2) Person + breach of biological expectations;
(3) Person + breach of psychological expectations;
(4) Animal + breach of physical expectations;
(5) Animal + breach of biological expectations;

(6) Animal + breach of psychological expectations;
(7) Plant + breach of physical expectations;
(8) Plant + breach of biological expectations;
(9) Plant + transfer of psychological expectations;
(10) Natural object + breach of physical expectations;
(11) Natural object + transfer of biological expectations;
(12) Natural object + transfer of psychological expectations;
(13) Artefact + breach of physical expectations;
(14) Artefact + transfer of biological expectations;
(15) Artefact + transfer of psychological expectations.

The template account is, on the whole, compatible with the anthropological record. It is no simple task to evaluate the relative frequency of templates. One serious difficulty is that the kind of evidence provided by anthropological reports, precious though it is, covers only one aspect of religious concepts, the overt, socially transmitted violations that make such concepts salient. A cognitive description of the templates also requires the tacit side, those background expectations that are used in producing inferences about religious objects and agents. Moreover, we have little or no reliable statistical information about the distribution of different types. However, we can be fairly confident about both ends of the distribution: concepts that we find virtually everywhere and concepts that are not reported so far in the anthropological record. This evidence confirms that most religious concepts are indeed based on one of the templates in this catalogue, given the ontological categories and principles normally developed by human minds. The evidence also suggests that the features that correspond to these templates are invariably *essential* to the representation of the religious concepts. In psychological terms, they are the *core features* that make such imagined objects and agents what they are. Religious and other supernatural concepts may or may not include other kinds of odd or unfamiliar or surprising information, but they invariably include a violation of ontological expectations. To explain why this is the case, we must turn to experimental evidence that highlights such processes in individual acquisition and memory.

Experimental evidence: recall and inference

We now have some experimental evidence that individual processes tend to favour supernatural concepts of the format described above. In a number of controlled studies we used artificial concepts that correspond to the templates described above, yet are not culturally familiar (see Barrett 1996; Barrett & Keil 1996; Boyer 1999; Boyer & Ramble n.d.). These studies measured recall for such concepts in the context of short narratives. Recall is particularly important because it is a necessary condition for cultural transmission. All else being equal, concepts that are recalled better than others have a higher potential for transmission. I will not describe here the specific protocols, materials or statistics of such studies but summarize their main findings.

(1) *Violations are recalled better than standard associations.* This occurs both in straightforward free recall tasks and in serial transmission studies, where the

PASCAL BOYER

participants' recall of a story is used as material for a second 'generation' of participants, and so on.[4]

(2) *Recall depends on ontological violations, not just on 'oddity'.* The previous results could be interpreted as the fact that violations are simply strange because they are unexpected. But further studies included items that are simply odd, unexpected but without ontological violations. Violations are better recalled and better transmitted than such 'odd' material. So strangeness cannot be the only explanation for recall of violations.

(3) *Violations do not contaminate people's intuitions.* Supernatural templates are procedures for handling information from intuitive ontology that produce salient concepts. This requires that intuitive ontology itself is not affected by salient concepts. This requires that intuitive ontology itself is not affected by the presence of such salient violations (otherwise they would not be salient). A good test of this would be whether we find situations in which people (i) represent a violation of intuitive expectations in the limited context of some religious notion and (ii) refuse similar violations in all other contexts. An illustration is Walker's (1992) study of transformations of natural kinds (an animal changes species) in both ritual and non-ritual contexts among the Yoruba of Nigeria. This showed that people who were willing to accept counter-intuitive transformations as ritual occurrences were equally confident that they could not happen in other contexts.

(4) *Inferences are governed by background default expectations.* Those intuitive expectations that are *not* violated are the main source of inferences about supernatural situations. This is confirmed by Barrett and Keil's ingenious experiments on God concepts (Barrett & Keil 1996). They elicited from the participants features that make God special. Subjects generally mention violations of theory-of-mind expectations, for example that God attends to everything at once. They then tested recall for stories that used these violations. They found that in the subjects' recall such features were generally replaced with more intuitive descriptions of cognitive functioning, taken from intuitive 'theory of mind'.

(5) *Violations with no inferential potential are not recalled.* Recall decreases when concepts violate intuitive expectations but do not allow default inferences for the category (e.g. a god that exists sporadically, a statue that is nowhere).

(6) *Sensitivity to violations is cross-culturally stable.* The God concept studies (see above, (4)) were replicated in India with similar results (Barrett 1998). Also, Boyer and Ramble (n.d.) tested recall for ordinary, 'odd' and violation concepts in two contrasted settings in Gabon and in Nepal. In the same way as in France and the United States, violations were recalled much better than either common associations or 'odd' material. Massive differences in cultural settings did not result in any significant differences in recall performance.

All this goes some way towards explaining why the supernatural concepts people are likely to entertain, acquire and store belong to the same 'catalogue' the world over. What specific 'choices' were made by the previous generations affects one's sensitivity to these concepts only marginally. There is a cross-cultural sensitivity to violations of intuitive expectations for ontological categories, which is not affected by the range of supernatural concepts used in the culture, by their variety or mode of transmission, or by people's com-

mitment to them. This may account for the first type of selection mentioned at the beginning, that which selects *some* products of human imagination to constitute the domain of culturally successful supernatural concepts.

Further selection in the supernatural imagination

Most supernatural concepts are found in popular fiction: folk tales, myths, anecdotes or urban myths. But some of these concepts are treated quite differently. They are construed as possibly describing real states of affairs. Also, they often trigger high emotional arousal or come to be linked to special kinds of experience. People may perform rituals about the entities described by these concepts. The notions are often associated with morality. They are connected to concepts of group identity, with important social consequences.

Such concepts (God, gods, ancestors, witches, spirits) are taken from the 'catalogue of the supernatural', and each of them belongs to one of the templates described above. But they are not just that. They seem to have important social implications. Indeed, this difference may well be the main basis for our intuitive understanding of the term 'religion'. That is, we tend to use that term when ontological assumptions about supernatural entities are combined with one or several of these features: commitment, moral understandings, group identity, ritual, special experience.

Considerable energy has been spent (or wasted) trying to turn this 'polythetic' characterization of religion into a more proper definition with necessary and sufficient conditions.[5] It is perhaps more useful to accept that these features are not always found together and to turn to questions of greater actuality. Are all supernatural concepts equally likely to be associated with such socially important effects? If not, why not? What gives a particular range of supernatural concepts a greater conceptual affinity to ritual, group identity, moral understandings or private experience?

In terms of our list of templates, it seems clear that such associations are made only with a specific subset of the supernatural templates, namely those that activate our 'theory of mind' expectations. This is true for instance of concepts of gods, ancestors and spirits, which correspond to:

(1) Person with breach of physical expectations;
(2) Person with breach of biological expectations.

In these cases 'intuitive theory of mind' is activated as the intuitive background associated with the category. In other cases, such as concepts of listening statues, activation of 'intuitive theory of mind' constitutes the violation part of the template:

(15) Artefact with transfer of psychological expectations.

It seems clear from the anthropological record that socially significant supernatural concepts are largely about agents spontaneously (and in large part

tacitly) represented as having psychological processes (perception, belief, inten-tion) that agree with our intuitive theory of mind.[6]

The idea that socially significant supernatural imagination is principally about imagined intentional agents is not altogether new, to say the least. But why should these concepts rather than others become associated with impor-tant social effects? Now nothing in what we said so far, in terms of concep-tual structure, would make it especially difficult to establish such associations with other templates for supernatural concepts. Stories that connect ritual, moral prescriptions or group identity to zombie-like animals or plants with strange physics would be equally salient, counter-intuitive, intriguing and the like. The privilege of intentional agency seems so obvious that we often forget that it requires an explanation that should be general (the phenomenon is found the world over) and psychological, as this is a matter of some concepts being easier to acquire and represent than others (Guthrie 1993).

Classical answers are of little help here. In anthropology we used to think, following Piaget, that projecting intentional agency onto non-intentional domains was a pervasive form of cognitive activity, especially prevalent in early childhood (Piaget 1954). This is not entirely plausible in the light of recent evidence concerning cognitive development. It seems that children in fact start with a strict demarcation between intentional agents and the rest (Bullock, Gelman & Baillargeon 1982; Rochat, Morgan & Carpenter 1997). Another explanation is that we generally tend to project human features onto unknown entities, because human features are most familiar. This is misleading on two counts. First, what is projected is not specifically human. Intuitive theory of mind is spontaneously activated by the presence not just of humans, but also of most animals. Second, not *all* human characteristics are projected, only intuitive theory of mind. In many religious concepts there is a projection of intentional agency but not of other human characteristics, such as having legs and arms, living in kin-based groups, eating cooked food.[7]

Even if we could explain why concepts of supernatural intentional agents are generally more salient than other supernatural concepts, this would not be enough. In any cultural group we find many concepts of supernatural agents, but only a small subset of these is associated with high commitment, emo-tional tenor or a connection to morality, ritual and group identity. There are more Mickey Mouse concepts than God concepts. The question is, what produces the difference between the 'serious' notions and the rest of the supernatural repertoire? We often treat this difference as a straightforward, unproblematic consequence of people's commitment. We assume that the main difference between Mickey Mouse and God, say for Christians, is that Chris-tians treat God as a real agent and Mickey Mouse as a fiction. True, but this is something we should *explain*, not take for granted. Over time, people are faced with many different versions of a particular religious concept and they themselves produce by inference many different variants of that concept. That some of these imagined agents are taken as more probably real than others is a function of these inferences, of the type of information that is associated with these different concepts.

So there must be some other feature that makes *some* psychology-based concepts special. To understand what that is, we must turn to another domain of evolved mental machinery, that of cooperation and information.

Mental mechanisms for social interaction

Let me start with a few commonplace remarks about general aspects of our evolutionary heritage:

(1) Humans depend more than any other species upon *information* about their environment. Most human behaviour is based on a rich and flexible data-base that gives parameters for action. Very little human behaviour can be explained or even described without taking into account the massive acqui-sition of information about surrounding situations. The proper ecological niche of humans is a 'cognitive niche' (Tooby & Cosmides 1990).

(2) Humans depend upon finely tuned *cooperation* with other human beings. Humans have for a long time (long enough to make a difference in evolu-tionary terms) lived in small groups and in intense social interaction. This interaction is not accomplished through stereotyped action sequences and in fact could not be. It requires subtle coordination between different people's actions.

(3) Because of these two facts, humans depend to an enormous degree on *information about other people's mental states*, in particular, what information they have and what their intentions are. No joint hunting expedition, war raid or marriage negotiation can be organized without precise monitoring of what other people want and believe.

That humans depend on cooperation creates all sorts of specific problems. Cooperation problems are *strategic* problems, where the value (the expected benefit) of a particular move depends on whether someone else makes a particular move (not necessarily the same one) (Schelling 1960). This creates 'commitment problems', that is, it becomes crucial to be able to estimate one's potential partner's willingness to cooperate or defect. A whole range of typically human characteristics is explained by these evolutionary factors, including:

(1) *A hypertrophy of social intelligence.* The human mind-design is character-ized by a hypertrophy of intuitive psychology. The set of intuitive principles that infer mental states — perceptions, beliefs and desires — from observed behaviour is vastly more complex than in any other species. Also, it appears very early and seems to engage dedicated brain structures. These con-stantly monitor what information is available to other agents and how they treat it.

(2) *A capacity to evaluate potential cooperators and to detect potential defectors.* For organisms that depend on cooperation, signals of trustworthiness in others are crucially important, as well as detection mechanisms for faked signals (Bacharach & Gambetta forthcoming). As evolutionary biologists have pointed out, the evolution of cooperation requires mental structures that allow (i) a precise identification of individuals, (ii) memories of past interaction and reli-ability of these individuals, (iii) some way of detecting cheating or likelihood of cheating, and (iv) some motivation to punish or exclude cheaters (Trivers 1985). Controlled studies confirm that we find such capacities in humans. For instance, a demonstration of the importance of cheater-detection is Cosmides's (1989) series of experiments on logical problem-solving. The subjects' per-formance in simple logical tasks is greatly increased when these are formu-lated as problems of cooperation and social contracts.

(3) *A repertoire of moral feelings.* Complex cooperation problems are to some extent solved by such feelings as an unmotivated disposition to honesty, an emotional preference for fair dealings, a disposition to feel deep anger at the mere suggestion of cheating, a disposition to punish defectors, a disposition to punish people who do not punish defectors. These feelings are not directly opportunistic. That is, in many situations they lead people to forgo the immediate benefits of relatively safe cheating or of profitable dealings with cheaters. However, such emotional dispositions help solve commitment problems in cooperation (Frank 1988). They create emotional states that are not voluntarily controlled and whose signals are difficult to fake. So emotional dispositions advertise reliability and deter defectors.

(4) *The stability of some un-fakeable (or difficult to fake) signals.* In humans and other species we find that evolution favours the development of some signals that are either impossible or at least difficult and costly to fake. In the case of humans, there is now good evidence that some clues concerning people's sincerity (tone of voice, facial expressions, gestures) are difficult to fake, indeed to a large degree outside voluntary control (Ekman 1985). People are often sensitive to these clues, even though they generally have little conscious access to the processes whereby a person's sincerity is gauged.

(5) *Easily produced self-deception.* The difficulty of efficient trickery (see above) is somewhat offset by another mechanism that is pervasive and extensively documented in experimental social psychology, and perhaps misleadingly labelled 'self-deception' (Trivers 1985). When it would be advantageous for agent A if agent B perceived a situation in a particular way, we find that A often tends to feel committed to this way of perceiving the situation. Controlled experiments show that people tend to adapt their way of seeing situations to the construal that it would be in their interest that others adopt. This is not self-deception in a hypocritical or manipulative sense, since such perceptions of situations are spontaneous and genuine. However, they constitute a strategic mechanism that allows sincere communication of a desirable perception.

(6) *Emotional rewards for gossip.* Humans spend a great deal of time acquiring information on other people. The ubiquity and importance of gossip suggest that there is an evolutionary disposition to pay attention to information concerning others, in particular in domains of great adaptive value such as sex, resources and status.[8] Because we live in groups and depend crucially on information, and because most of the crucial information is information about other people's actions and intentions, we are predisposed to enjoy communicating about third-party actions and intentions, although the information conveyed may be of no immediate benefit. This also explains why gossip is as universally despised as it is enjoyed. It undermines our great efforts at impression management and threatens our control over the information we want to transmit about our own actions and intentions.

This description of specialized mental machinery explains how people assess particular situations and use evolved dispositions to help them make roughly adequate choices given the circumstances. Social psychologists have observed these effects in many different domains of social interaction. Controlled studies show that intuitions about people's attractiveness, friendliness, trustworthiness, honesty ' the like, are the outcome of exceedingly complex computations

on the basis of subtle signals (Bacharach & Gambetta forthcoming). But these computational processes remain largely inaccessible to conscious access. What is accessible to consciousness is typically a feeling that a certain person is attractive, reliable and so forth. We track information about other agents' moves and about the information they have access to, but all this is pursued effortlessly, without conscious deliberation.

The cognitive devices that handle such computations are *information-processing* devices, whose input consists in information that is potentially relevant to interaction: facial cues for emotions, reputation of the potential cooperator, past episodes of interaction with him or her, social categories to which he or she belongs, contingent signs of trustworthiness and the like. This information is then handled differently according to the situation at hand as different contexts (e.g. kinship relations, social exchange, mate-choice) probably require different principles. The output is a set of particular intuitions about people, about situations, about the emotional tenor of situations, as well as preferences between possible outcomes or between plans for future action.

Strategic information and varieties of agents

Let us call *strategic information* the domain of information that feeds into information-processing devices geared to regulating social interaction. In many species, objective features differentiate strategic information from other information: special signals, special contexts and the like.[9] In humans, by contrast, *any* aspect of a situation potentially contains strategic information. Whether it does or not depends on a representation of the particular situation at hand. Depending on how I represent the situation, that you have meat in your refrigerator may be non-strategic to me (in most cases) or strategic (if meat was stolen from my pantry, or if I am hungry, or if you always declared you were a vegetarian). That you went to the next village yesterday may be non-strategic (if all I infer is that you were away) or strategic (if I suspect that you went there to meet a potential sexual partner). That you talked with So-and-so may become strategic if I suspect that the two of you are involved in some plot against me or a potential coalition with me. The distinction between strategic and non-strategic depends on a representation of the particular situation and can be expressed in functional terms:

> Strategic information is the subset of all the information currently available (to a particular agent, about a particular situation) that activates mental capacities that regulate cooperation.

Note that this does not in any way entail that strategic information is more important than other information. To know about our coworkers' sexual peccadilloes (strategic) is in most circumstances of negligible import; to know how to flee from predators (non-strategic) is much more useful. Also, this definition does not imply that one is right to consider some information as strategic. The distinction between strategic and non-strategic information relative to a situation, is all in the eye of the beholder, who may well be wrong.

Humans not only represent strategic information, they also represent the extent to which other people have strategic information. For instance, given a particular situation where you have something that I want, I automatically form a representation, not just of the fact that you have that which I want, but also that you may be aware of the fact that I want it, and that this may have some influence on your intentions. Such complex inferences are supported by our theory of mind, which represents other people's mental states and their access to information.

Now one fundamental principle of our intuitive theory of mind is that access to information is *imperfect*. Given a situation, and given some information about that situation, we do not automatically presume that this information is equally accessible to everyone. For instance, if I remove your keys from your pocket while you left the room, I expect that you will not be aware of what I did. I expect that you will be surprised when you cannot find your keys. Normal children from the age of four routinely solve experimental tasks that require evaluating such obstacles to information transfer (Perner, Leekam & Wimmer 1987; Wimmer & Perner 1983). This principle of imperfect information is so fundamental that not having it in one's cognitive equipment results in pathologies like autism (Baron-Cohen 1995; Baron-Cohen, Leslie & Frith 1985; Leslie & Frith 1987). The principle applies to information in general and therefore to the subset of information that is strategic information:

> Social interaction is informed by the tacit assumption that people (ego and others) are *mere strategic agents*, that is, agents whose access to strategic information is neither perfect nor automatic.[10]

That is, given a particular situation and some information about it that is strategic to you (that is, activates your mental capacities for social interaction), you cannot automatically *presume* that other people, in particular other people involved in the situation, also have access to that information. You went to the other village last night for a secret rendezvous. It is clear to you that the identity of the person you met is strategic information; but it is not clear to what extent that information is available to other people. Indeed, you may hope that it is not (for fear of scandal) or wish that it were (so you can brag about the episode). Humans generally spend a great deal of time and energy wondering whether other people have access to some information that is strategic in our view, wondering what inferences, intentions or plans these other people draw from that information, trying to control their access to such information and trying to monitor and influence their inferences on the basis of such information. All these complex calculations are based on the assumption that other agents' access to strategic information is complex and generally imperfect.

Supernatural agents of a special kind

The point of all this is that we can now describe in simple yet precise terms

> *Full-access strategic agents* are (imagined) agents who we presume have access to any piece of information that is strategic.

That is, given a particular situation, and given some information that one represents as strategic about that situation, one assumes that the 'full-access strategic agent' has access to that information. We do not need to encounter such agents to represent them. Indeed, all this fantasy requires is that in some circumstances there are no obstacles between the imagined agent and relevant strategic information.

This is of great consequence for our understanding of supernatural concepts. So far, I have described supernatural agents in terms of their general features and powers, in particular in terms of what violations of intuitive expectations are included in their description. But one should also ask, What are the *contents* of the thoughts and intentions attributed to imagined agents? In other words, what do people assume spirits and gods think about?[11]

Imagined religious agencies are in general credited with good access to information. That they appear at several places at the same time or become invisible gives them the means to hold information that real agents have more difficulty acquiring. I do not want to suggest that such agents are always considered to be wiser than mere mortals. Actually, we know of many cases where they are represented as intrinsically stupid. So the point is not that they know *better* but simply that they often seem to know *more*. Indeed, in the many narratives that include such agents as well as human ones, the scenarios in which a religious agent has information that a human agent does not possess greatly outnumber descriptions of the converse situation. God knows more than we know, the ancestors are watching us. This can be summed up in more precise terms as follows:

> Some supernatural agents (i.e. agents with counter-intuitive features as defined above) are tacitly represented as full-access strategic agents.

That is, people tacitly construe them as always having access to whatever information they themselves represent as relevant to social interaction. Consider this: In most local descriptions of spirits and other such agents, we find the assumption that they have access to information that is not available to ordinary folk. But this, in general, turns out to be information about people's motivations, about their intentions to harm or help other people, about their actions and so on.

An interesting limiting case is the concept of gods who know *everything*. The theological, literate version of such concepts stipulates that the god has access to all information about the world from all possible angles. But we know that people's actual concepts often diverge from theological understandings, as Barrett and Keil (1996) demonstrated, so we may wonder whether people actually represent omniscient gods in this way. If they did, they would assume that all pieces of information about all aspects of the world are equally likely to be represented by God. In this case the questions, Does God know the contents of every refrigerator in the world? Does God perceive the state of every machine in operation? Does God know what every single insect in the world is up to? would be every bit as natural as, Does God know that

you are lying? Does God know that I misbehaved? But that is not the case. The first set of questions seem strange; answering them requires more deliberation than answering the second set. In people's representation, it seems that God is assumed to know primarily of morally relevant aspects of human situations. That is why God may in fact be thought to represent the contents of your refrigerator (if that includes items you stole from your neighbours), the state of some machines (if you use them to harm people) and of insects (if they are a plague we wished upon the enemy). In such situations that information is strategic. Intuitively, people who represent such situations immediately assume that God represents the information that is strategic to them.

More generally, religious agents are explicitly described as having access to a lot of information and tacitly presumed to have full access to strategic information. Obviously, such imagined agents may be represented as having access to other, non-strategic information as well but, I would claim, it is difficult to find concepts of religious agents without access to strategic information. In general, what religious agents know is what matters to social interaction in a human group. Note that this is a matter of spontaneous inference. This quality of imagined agents does not need to be explicit any more than the assumption 'solid objects fall downwards' needs to be entertained explicitly. But the presumption of full access to strategic information does inform people's inferences about certain particular types of supernatural agents.

Implications of full-access agent concepts

The question of why some concepts of supernatural agents are associated with important social effects (the God vs. Mickey Mouse question) may be much easier to approach if we accept that:

> Concepts of supernatural agents tacitly construed as full-access strategic agents are more likely than other concepts of supernatural agents to be associated with special experience or morality or group identity or ritual.

In a given situation, if a subset of information about that situation is relevant to strategic mechanisms, and if people imagine supernatural agents, we will often find that *some* of these agents are tacitly presumed to have access to the strategic information, are what I called full-access strategic agents. This could make sense of the variety of features that make up the family resemblance of 'religious' concepts. In this article I cannot do more than outline these connections.

First, the tacit assumption affords a straightforward connection between supernatural agents and moral understandings. We find this connection in many human groups, although the way in which it is construed is highly variable. In some places it is a theological connection (the gods laid down the rules we live by). In most human groups there is no such theology, but it just goes without saying that ancestors and spirits and other such agents are concerned with the way people behave and use their powers against those who violate moral prescriptions. We often assume that this connection results fr   ltural axioms or that morally relevant supernatural agents are some-

how useful for social cohesion. But there is a simpler account. Moral rules are easily acquired by human minds because of a host of computational predispositions, as a consequence of our dependence upon others' cooperation. This also makes the tracking of strategic information held by other people a constant cognitive activity. So what seem to be two different cultural domains – moral prescriptions on the one hand, concepts of full-access supernatural agents on the other – activate the same inferential processes, geared to establishing cooperation and evaluating the information that motivates other people's moves.

Second, the connection between religious agency and group identity, though it is far too complex to examine here, may be better understood in the light of the present description of strategic agents. Mental mechanisms that regulate cooperation are involved in evaluating a realistic level of trust, given (among other cues) genetic and ethnic distance between partners. If concepts of strategic agents are (in part) signals of reliability, we should predict, not necessarily that people will trust potential partners who share their religious commitments, but that they should distrust people who do not.

Third, at least some of the emotional arousal that is sometimes connected to religious agents makes more sense in the context of strategic agency connected to moral understandings and cooperation. For instance, we observe that people often pay considerable attention to (and sometimes resent) the fact that others are not committed to particular beliefs. This may be less surprising if we recall that adaptations for cooperation are themselves emotional mechanisms, and that clues (however indirect) of non-cooperation have notable emotional effects. The particular emotional tenor of religious activity may be diverse and variously glossed as trust or misgivings or hope. These diverse emotions make sense when directed at (or putatively caused by) agents who hold information essential to decision-making. Emotions connect directly to planning and social interaction. The assumption that strategic information is actually represented by some agent with full access changes the tenor of interaction, makes certain moves more or less possible or beneficial, and this is reflected in emotional states.

We anthropologists often imply that people first have concepts of powerful supernatural agents and then connect such concepts to socially important phenomena like morality, ritual and group identity because of that alleged power. This is also how many people explicitly describe their religious notions. But cognitive scientists know that such explicit accounts of one's own inferences are often based on *post hoc* rationalization. Indeed, in this case things seem more likely to go the other way around. We have evidence that specialized mental mechanisms track cues for cooperation, and highlight information that is relevant to cooperation (strategic information). We also have evidence that these mental mechanisms are tightly connected to moral feelings and to group identity. The present proposal is that some products of supernatural imagination are particularly relevant to these mental mechanisms, and this is consistent with anthropological evidence concerning what people think gods and spirits know. So imagined agents may not be connected to social interaction because they are powerful; they may be represented as powerful because they were tacitly represented in a way that connects them to crucial dimensions of social interaction.

Relevance of full-access agency concepts

Assuming that all these features of religious agency may be *consequences* of postulating full-access strategic agents, it remains to explain why people represent such agents. Why would people assume that something in the world, over and above the real agents they deal with, has strategic information? But the question is misguided, at least in this formulation. The fact that some kinds of notions are selected in cultural transmission does not depend on people's inclinations but rather on the aggregation of acquisition and memory processes over which they have no control. So a better formulation would be, Why are accounts of such strategically informed imaginary 'somethings' more likely to be acquired and transmitted than other possible accounts?

This, as in the first selection process described in the first part of this lecture, is primarily a matter of differential cognitive effects. Some cultural representations are *more relevant* than others, in the cognitive sense of producing more inferences with lesser cognitive effort. So a concept of full-access strategic agent would be relevant to the extent that (1) it is couched in a format that can be handled by some mental mechanism; (2) it allows *more* inferences than other concepts, when combined with stored information in that mechanism, or requires *less* processing to generate inferences.[12]

Concepts of full-access strategic agents are not costly in terms of processing. What distinguishes full-access from mere strategic agents is that, when we represent the former, we do not represent what obstacles lie between them and strategic information, as we do when we estimate other people's access to information. So religious agents are ordinary agents *minus* a crucial feature. To represent them, people run routine theory of mind inferences, but in a simpler way. The cognitive effort of representing ancestors fully apprised of our misdemeanours may actually be less than that of representing actual people who are only partly informed.

These concepts are also rich in inferential potential; that is, richer than equivalent supernatural concepts without the full-access assumption. That a full-access agent is imagined to be around changes the value and possible outcome of possible courses of events and of intended courses of action. In most situations of social interaction, we need access to other people's actions and intentions, but we also need to protect ourselves by broadcasting only a certain description of our own intentions and actions. A strategic agent typically sees through all this and has access to real actions and intentions, rather than the public version. So imagining such agents creates a distinct representation of any situation. Note that this description from an imagined resentation of any situation. Note that this description from an imagined agent's viewpoint may be completely wrong. What matters here is not that it is true but that it is *richer* in inferences than the description yielded by other supernatural concepts.

These somewhat speculative considerations could be summed up as the following principle of selection:

Cultural material that includes concepts of imagined agents is made more relevant by the tacit assumption that these agents have full access to strategic information.

Producing more effects at a lesser cost would ensure that such material is

domains. This, obviously, would also explain why we find such an advantage for supernatural concepts that activate intuitive psychology. The least costly (and inferentially richest) way of representing something in the world that has access to *any* information is to apply to it our intuitive psychology. If there is something that has strategic information, the story that this something has a *mind* is the simplest one available.

Functional origins of religion

What is the origin of religion? We find a plethora of functional answers in anthropology and the philosophy of religion: to explain the natural world, to justify a social order, to account for our mortality, to make the world meaningful and the like. As we teach our students, none of these answers is really satisfactory. They are all untestable, as they generally refer to historical scenarios for which there is no conceivable evidence. Also, they are often ethnocentric, describing cognitive and emotional aspects of religion that are not really general of human societies. Finally, they predict a whole variety of possible concepts that could fulfil the functions described. But we only find a small subset of those in the anthropological record.

Progress in our knowledge of evolved human capacities provides a better account that is testable, based on what we find in the anthropological record rather than in familiar settings, and that predicts only the narrow range of concepts observed rather than a variety of other possible concepts. Religious concepts are not around because they are good for people or for society or because of an inherent need or desire to have them. They are around because they are more likely to be acquired than other variants. A benefit of this functional account is to explain why we find *these* religious concepts and not others. There is a cost, too, since we have to provide independent evidence for the cognitive processes involved in cultural selection. We have some such evidence already. Much remains to be done, but progress in our understanding of the mind is gradually transforming many mysteries of culture into mere problems, including that of the origins of religion.

NOTES

[1] Throughout this lecture I assume that such a naturalistic account is both possible and desirable. The research on religion presented here was initially inspired and strongly influenced by general arguments to that effect (Sperber 1985; 1996) and by their application to specific cultural domains (Atran 1990; Atran & Medin 1999; Hirschfeld 1996). The evolutionary background to this approach is outlined in Tooby & Cosmides (1992), some anthropological consequences in Bloch (1998) and Boyer (1998).

[2] This section summarizes arguments presented in greater detail in Boyer (1994; 1998).

⁵I cannot here give detailed references for all these aspects of intuitive ontology. A good presentation of the relevant psychological research can be found in Hirschfeld & Gelman (1994) and Sperber, Premack & James-Premack (1995). For general consequences on cultural transmission, see Boyer (1998).

⁴This in effect is a more tightly controlled version of Bartlett's (1932) classic experiments on memory for folk-tales.

⁵See Needham (1975) on the fact that most anthropological concepts are 'polythetic' and on the futility of trying to turn them into proper analytic categories.

⁶A good illustration of this preference for templates with activation of intuitive psychology is the difference in frequency between concepts of 'spirits' and of 'zombies'. These constitute symmetrical violations. A spirit is a person with standard psychological properties and special physical ones. A zombie is a person with standard physical properties but special psychological ones. Now the spirit type of religious concepts is much more frequent than the zombie type. Further, wherever people have a concept of zombie, they invariably complement that with the notion of some spirit or witch or god who 'remote-controls' the zombie (Boyer 1996).

⁷Whenever we encounter concepts that actually include these other features of human life, they invariably include mental processes as well. The reverse is not true (Boyer 1996). So it really is the projection of psychological processes that matters here.

⁸Surprisingly, there is very little anthropological evidence and theory on this fundamental activity. I am inspired here by two remarkable exceptions, Haviland (1977) and Gambetta (1994).

⁹The situation is very different in most other species, even in closely related ones. Among chimpanzees, for instance, specific signals indicate a willingness to engage in peaceful cooperation (grooming) or to challenge some other individual's status (shaking branches and emitting special cries). In monkeys, alarm cries are not just segregated from other voicings but specialized for different kinds of hazards.

¹⁰Note that we also sometimes interact with agents that are not really presumed to have much access to strategic information at all, like infants or animals. These could be called 'infra-strategic' agents. They are agents and they have access to some information, but we do not spontaneously assume that they represent the strategic aspects of a situation or produce inferences on the basis of that strategic information.

¹¹Remember that we are talking here about people's tacit presumptions. People may have such presumptions without being aware of them. Also, such tacit presumptions may or may not be officially expressed in or sanctioned by a theology.

¹²This refers to a technical description of relevance in terms of relative processing cost and relative effects which can be operationalized (see Sperber & Wilson 1995). To illustrate this with low-level process: Faces are more relevant than legs in distinguishing between people, not because they are intrinsically more distinctive, but because (1) perceptual routines specialized in face-recognition translate faces into distinctive patterns, (2) this allows a comparison with stored patterns, which (3) quickly delivers recognition decisions. Information that is not couched in a format accessible by a mental mechanism (e.g. for humans, facial traits in giraffes) is not relevant.

REFERENCES

Atran, S. 1990. *Cognitive foundations of natural history: towards an anthropology of science*. Cambridge: University Press.

—— & D. Medin (eds) 1999. *Folkbiology*. Cambridge, Mass: M.I.T. Press.

Bacharach, M. & D. Gambetta forthcoming. Trust in signs, in *Trust and social structure* (ed.) K. Cook. New York: Russell Sage Foundation.

Baron-Cohen, S. 1995. *Mindblindness: an essay on autism and theory of mind*. Cambridge, Mass: M.I.T. Press.

—— A. Leslie & U. Frith 1985. Does the autistic child have a 'theory of mind'? *Cognition* 21, 37–46.

Barrett, J.L. 1996. Anthropomorphism, intentional agents, and conceptualizing God. Ph.D. dissertation, Cornell University.

—— 1998. Cognitive constraints on Hindu concepts of the divine. *Journal for the Scientific Study of Religion* 37, 608–19.

—— & F.C. Keil 1996. Conceptualizing a non-natural entity: anthropomorphism in God concepts. *Cognitive Psychology* 31, 219–47.

Bartlett, F.C. 1932. *Remembering*. Cambridge: University Press.

Bloch, M. 1998. *How we think they think: anthropological approaches to cognition, memory and literacy*. Boulder, Col.: Westview Press.

Boyer, P. 1994. *The naturalness of religious ideas: a cognitive theory of religion*. Los Angeles: University of California Press.

—— 1996. What makes anthropomorphism natural: intuitive ontology and cultural representations. *Journal of the Royal Anthropological Institute* (N.S.) 2, 1–15.

—— 1998. Cognitive tracks of cultural inheritance: how evolved intuitive ontology governs cultural transmission. *American Anthropologist* 100, 876–89.

—— 1999. Human cognition and evolution. In *Anthropological theory today* (ed.) H. Moore, 206–33. Cambridge: Polity Press.

—— & C. Ramble n.d. Cognitive templates for religious concepts: cross-cultural evidence for recall of counter-intuitive representations. MS.

Bullock, M., R. Gelman, & R. Baillargeon 1982. The development of causal reasoning. In *The developmental psychology of time* (ed.) W.J. Friedman, 209–54, New York: Academic Press.

Cosmides, L. 1989. The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* 31, 187–276.

Ekman, P. 1985. *Telling lies*. New York: Norton.

Frank, R. 1988. *Passions within reason: the strategic role of the emotions*. New York: Norton.

Gambetta, D. 1994. Godfather's gossip. *European Journal of Sociology* 35, 199–223.

Guthrie, S.E. 1993. *Faces in the clouds: a new theory of religion*. New York: Oxford University Press.

Haviland, J.B. 1977. *Gossip, reputation, and knowledge in Zinacantan*. Chicago: University Press.

Hirschfeld, L.A. 1996. *Race in the making: cognition, culture and the child's construction of human kinds*. Cambridge, Mass: M.I.T. Press.

—— & S.A. Gelman (eds) 1994. *Mapping the mind: domain specificity in culture and cognition*. New York: Cambridge University Press.

Leslie, A. & U. Frith 1987. Meta-representation and autism: how not to lose one's marbles. *Cognition* 27, 281–94.

Needham, R. 1975. Polythetic classification: convergence and consequences. *Man* 10, 349–69.

Perner, J., S.R. Leekam, & H. Wimmer 1987. Three year olds' difficulty with false belief. *British Journal of Developmental Psychology* 5, 125–37.

Piaget, J. 1954. *The child's construction of reality*. New York: Basic Books.

Rochat, P., R. Morgan, & M. Carpenter 1997. Young infants' sensitivity to movement information specifying social causality. *Cognitive Development* 12, 441–65.

Schelling, T. 1960. *The strategy of conflict*. Cambridge, Mass: Harvard University Press.

Sperber, D. 1985. Anthropology and psychology: towards an epidemiology of representations. *Man* 20, 73–89.

—— 1996. *Explaining culture*. Oxford: Blackwell.

—— D. Premack, & A. James-Premack (eds) 1995. *Causal cognition: a multidisciplinary debate*. Oxford: Clarendon Press.

—— & D. Wilson 1995. *Relevance: communication and cognition*. (2nd edition). Oxford: Blackwell.

Tomasello, M., A.C. Kruger, & H.H. Ratner 1993. Cultural learning. *Behavioral and Brain Sciences* 16, 495–510.

Tooby, J. & L. Cosmides 1990. The past explains the present: emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology* 11, 375–424.

—— & —— 1992. The psychological foundations of culture. In *The adapted mind: evolutionary psychology and the generation of culture* (eds) J. Barkow, L. Cosmides, & J. Tooby, 19–136. New York: Oxford University Press.

Trivers, R. 1985. *Social evolution*. Menlo Park, Cal.: Cummings.

Walker, S.J. 1992. Supernatural beliefs, natural kinds, and conceptual structure. *Memory and Cognition* 20, 655–62.

Wimmer, H. & J. Perner 1983. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–28.

Origines fonctionnelles des concepts religieux: sélection ontologique et religieuse par les esprits évolués

Résumé

Les concepts religieux qui ont du succès culturellement résultent de sélections qui accroissent la probabilité que certains concepts, plus que d'autres, soient acquis, conservés et transmis facilement. Parmi les constructions de l'imagination humaine, certaines sont reliées à des principes ontologiques intuitifs de telle sorte qu'elles constituent un petit catalogue de concepts surnaturels qui ont du succès culturellement. Des preuves expérimentales et anthropologiques apportent confirmation de la salience et du potentiel de transmission de ce catalogue. Parmi ces concepts surnaturels, les aptitudes cognitives à l'interaction sociale introduisent une sélection supplémentaire. En conséquence, certains concepts d'agents surnaturels sont associés à la moralité, à l'identité de groupe, au rituel et à l'émotion. Ces agents surnaturels 'religieux' typiques sont tacitement présumés avoir accès à l'information qui est cruciale pour l'interaction sociale, une supposition qui renforce leur diffusion dans les groupes humains.

Institut des Sciences de l'Homme, 14 avenue Berthelot, 69363 Lyons, France. pboyer@ish-lyon.cnrs.fr