ELSEVIER

Cognitive Science 25 (2001) 535–564

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Cognitive templates for religious concepts: cross-cultural evidence for recall of counter-intuitive representations

Pascal Boyer[a,*], Charles Ramble[b]

[a]*Washington University One, Brookings Drive, St. Louis, MO 63130, USA*
[b]*Wolfson College, Oxford University, Linton Rd, Oxford OX2 6UD, UK*

Abstract

Presents results of free-recall experiments conducted in France, Gabon and Nepal, to test predictions of a cognitive model of religious concepts. The world over, these concepts include violations of conceptual expectations at the level of domain knowledge (e.g., about 'animal' or 'artifact' or 'person') rather than at the basic level. In five studies we used narratives to test the hypothesis that domain-level violations are recalled better than other conceptual associations. These studies used material constructed in the same way as religious concepts, but not used in religions familiar to the subjects. Experiments 1 and 2 confirmed a distinctiveness effect for such material. Experiment 3 shows that recall also depends on the possibility to generate inferences from violations of domain expectations. Replications in Gabon (Exp. 4) and Nepal (Exp. 5) showed that recall for domain-level violations is better than for violations of basic-level expectations. Overall sensitivity to violations is similar in different cultures and produces similar recall effects, despite differences in commitment to religious belief, in the range of local religious concepts or in their mode of transmission. However, differences between Gabon and Nepal results suggest that familiarity with some types of domain-level violations may paradoxically make other types more salient. These results suggest that recall effects may account for the recurrent features found in religious concepts from different cultures. © 2001 Cognitive Science Society, Inc. All rights reserved.

Keywords: Religion; Concepts; Recall; Distinctiveness

1. Introduction

Religious concepts label and describe supernatural agencies in ways that are specific to each culture. The concepts may describe a unique, omniscient and omnipotent God, as in

* Corresponding author. Tel.: +1-314-935-5252; fax: +1-314-935-8535.
E-mail address: pboyer@artsci.wustl.edu (P. Boyer).

0364-0213/01/$ – see front matter © 2001 Cognitive Science Society, Inc. All rights reserved.
PII: S0364-0213(01)00045-3

Judaism, Christianity and Islam, or several gods with their particular features, or a variety of ghosts, spirits, ghouls, zombies, fairies, djinns, as well as specific artifacts (e.g., statues that listen to people's prayers), animals (e.g., jaguars that live in the sky) or parts of the natural environment (e.g., mountains that think, rivers that protect people, etc.) (see e.g., Child & Child, 1993). This variety might seem to suggest that "anything goes" in this domain. Anthropologists have long suspected that there were in fact limits to variability. However, there was no cognitive account of the processes that would make certain types of concepts more "culturally successful" than others.

Cognitive studies of religion start from the premise that religious concepts are governed by the same kind of constraints as other concepts and can be investigated in the same way (Goldman, 1964; Watts & Williams, 1988; Barrett, 2000). A number of anthropologists have argued that religious concepts do not in fact constitute an autonomous "domain" (Spiro & D'Andrade, 1958; Sperber, 1985; Lawson & McCauley, 1990; Bloch, 1992; Whitehouse, 1992; Guthrie, 1993; Dulaney & Fiske, 1994; Barrett & Keil, 1996). Information derived from nonreligious conceptual schemata constrains religious ontology (Guthrie, 1993; Barrett & Keil, 1996), ritual taboos (Dulaney & Fiske, 1994), concepts of ritual action (Lawson & McCauley, 1990; Houseman & Severi, 1998), the modes of transmission of religion (Whitehouse, 1992) as well as some developmental aspects of religious belief (Boyer & Walker, 1999).

The experiments presented here test the predictions of a particular cognitive account of religious concepts (Boyer, 1992; 1994, see also Sperber, 1996; Barrett, 1996). This account is based on a distinction between information represented at the level of (roughly) "basic" kinds (henceforth "kind-concepts") and that associated with broader ontological categories ("domain-concepts"), such as person, artifact, animal, inanimate natural object, plant, and so forth (see Hirschfeld & Gelman, 1994 for a survey of these categories). Whenever an object is identified as belonging to one such domain, this triggers specific, principled expectations that go beyond the information given and establish causal links between observed features and underlying structure. Domain concepts are described by developmental psychologists as based on "skeletal principles" (Gelman, 1990), more abstract "modes of construal" that emphasize particular forms of causation (Keil, 1986; 1994) or "foundational theories" that specify specific principles for each domain (Wellmann & Gelman, 1992). This domain-level information produces expectations intuitively applied to new objects, however unfamiliar, if they are identified as members of a particular ontological domain (Keil 1979).[1]

The distinction between kind- and domain-concepts is relevant here because most religious notions imply a particular treatment of information associated with domain-concepts:

1. Religious concepts generally include explicit *violations of expectations* associated with domain-concepts. For instance, spirits and ghosts go through physical obstacles or move instantaneously, thereby violating early developed expectations about solid objects (Spelke, 1988). Eternal gods, metamorphoses, chimeras and virgin birth go against entrenched expectations about living things (see e.g., Carey, 1985; Walker, 1992; Keil, 1994). Gods that perceive everything or predict the future violate early-developed assumptions about intentional agency and about the causal links between events, perceptions and beliefs (Wellmann, 1990; Perner, 1991; Gopnik, 1993). An-

thropomorphic artifacts with a psychology, like for instance "listening statues" that hear and understand people's prayers, are construed as artifacts with nonstandard properties, intentional properties in this case (Guthrie, 1993; Boyer, 1996b; Barrett, 2000).

2. Religious concepts invariably require that relevant *nonviolated assumptions are tacitly activated* by default. For instance, while people represent spirits as physically nonstandard agents, they also tacitly represent them as cognitively standard agents, spontaneously extending intuitive psychological expectations to supernatural agents (Barrett & Keil, 1996). In a symmetrical way, "zombies" are construed as cognitively nonstandard, but "intuitive physics" is tacitly applied to them by default; their bodies are supposed to obey the same physical constraints as other solid objects (Boyer, 1996b).

This model predicts that culturally successful religious concepts belong to a small number of recurrent types or *templates*. A template has the following entries:

[1] a pointer to a particular domain concept
[2] an explicit representation of a violation of intuitive expectations, either:

[2a] a *breach* of relevant expectations for the category, or
[2b] a *transfer* of expectations associated with another category;
[3] a link to (nonviolated) default expectations for the category.

For instance, most concepts of spirits, ghosts and ancestors correspond to a particular template, where [1] is a pointer to the category person, [2a] is the assumption that these special persons have counterintuitive physical properties (e.g., they can go through walls, violating "intuitive physics" for solid objects) and [3] specifies that these agents confirm all intuitive "theory of mind" expectations. These three features are found in most concepts of ghosts or spirits or gods, which is why we will say that they correspond to a single template. By contrast, concepts of statues that listen to people's prayers correspond to a different template, where [1] specifies the category artifact, [2b] mentions a transfer of intentional properties to these artifacts, and [3] confirms that ordinary physical properties of artifacts are still relevant.

Religious *concepts* are more specific than *templates*, in that they add to the template two other entries:

[4] a slot for additional encyclopedic information;
[5] a lexical label.

The Western concept of 'ghost', for instance, implies the template described above, plus [4] information like "ghosts often come back to where they used to live" and [5] "ghost." The place-holder [4] is where all sorts of culturally specific information can be inserted, making the religious *concepts* of different places obviously different even though they may correspond to the same *templates*.

In this view, the reason why the 'ghost' concept of a particular culture is easily transmitted lies not in the particulars of this concept but in the template it shares with other concepts of that type, found in many cultures. Anthropological evidence seems to support this prediction,

in that limited violations of domain-level expectations constitute a narrow "catalogue" of templates that accounts for most kinds of supernatural concepts. The concepts found in the anthropological record generally include salient violations of intuitive expectations in their explicit characteristics; they also include tacit activation of nonviolated assumptions for the relevant domain-concepts (Boyer, 1996a; Boyer, 1996b).

The anthropological model suggests, first, that violations of intuitive expectations for domain-concepts are probably more salient than other types of cultural information, thereby leading to enhanced acquisition, representation and communication; second, that most inferences about religious concepts are driven by nonviolated, mostly tacit domain-level expectations, while explicit, consciously accessible, officially sanctioned theologies have little effect on representation and inference.

This latter point was illustrated by Barrett & Keil's work on God-concepts (Barrett & Keil, 1996). This showed that even the most central aspects of a religious belief (e.g., for Christians, that God can attend to several events at once) are by-passed in favor of a more anthropomorphic concept (God like other agents attends to situations serially) when task-demands require fast activation of an "on-line" God-concept. In story recall, participants produce inferences on the basis of their intuitive expectations about psychological functioning, even when these inferences contradict their official, reflective beliefs, an effect that was replicated in India (Barrett, 1998). This would suggest that theologies have limited effects on concepts, compared with the inferential potential provided by domain-level expectations. This could explain a phenomenon familiar to most anthropologists: Although some theological systems do not correspond to the violation+default expectations model described above, their more widespread, "popular" interpretations generally distort the theological doctrine towards one of the templates described here (Boyer 1992).

The point of the present exploratory studies was to examine the effects of the *violation* part of the templates. We focused on the possible contribution of *recall* to the cultural fitness of "counterintuitive" concepts. Obviously, recall is only one of the conditions of cultural spread. Its effect interacts with the initial frequency of a particular cultural input, attention paid to that input, ease of reproduction, motivation for communication, conformism effects, and so forth to produce general cultural stability. We cannot limit what is represented to what is recalled. However, recall for particular input and distortions of cultural input created by individual recall are of prime interest because better recall is a condition of greater diffusion within a cultural environment (Sperber, 1985; Sperber, 1991). That is, considering opposite extremes for the sake of argument, we can expect, all else being equal, concepts that are very easy to recall to spread in a cultural environment and concepts that are intrinsically difficult to recall to spread less.

Surprisingly, there are few studies of recall for cultural material and their conclusions are ambiguous. On the basis of suggestive studies, Bartlett argued that recall is a "constructive" process that reframes exotic material in terms of familiar "schemata" (Bartlett, 1932; Bergman & Roediger, 1999). However, Bartlett's design made it difficult to set apart the contribution of "exotic" (nonculturally familiar) elements from that of expectation-violations. Higher recall for incongruous or surprising material is a familiar result from studies of story-recall (Hudson, 1988; McCabe & Peterson, 1990; Davidson, 1994), of "memory for scripts" (Graesser, Woll, Kowalski, & Smith, 1980; Brewer, 1985), of naturally "scripted"

situations (Nakamura et al., 1985), and of expectancy-incongruent material in the social domain (Stangor & McMillan, 1992). However, in all these studies the elements that go against conceptual expectations also disrupt script-structure. In a typical religious narrative by contrast, counterintuitive elements are neither peripheral nor disruptive. They constitute the focal points of the stories or events.

Semantic incongruity is also investigated in list designs, avoiding the possible confounds created by a narrative format. Such studies have generally shown classical "von Restorff" effects (von Restorff 1933; Hunt, 1995). That is, in the same way as perceptually distinctive stimuli are recalled better than other ones, conceptually incongruous material is retrieved better than common conceptual associations in recall tasks (McDaniel & Einstein, 1986; Schmidt, 1991; Waddill & McDaniel, 1998). The effect is probably not caused by differential encoding, since incongruous items that appear early in a list also produce the effect (Hunt, 1995) and it does not always require the generation of bizarre mental imagery (Worthen, 1997). Incongruous items enjoy an advantage at the representation or retrieval stage (Waddill & McDaniel, 1998) which may depend upon two different processes: an intralist comparison that makes such items "distinctive" and a violation of expected associations with active schemas (Worthen et al., 1998), although the specific process whereby incongruous information is compared to active schemata is not really elucidated (see Waddill & McDaniel, 1998, pp. 117–119). The empirical evidence accumulated over decades shows that intralist distinctiveness is a robust method for assessing the effects of incongruous conceptual associations. However, studies of distinctiveness have generally used violations of *kind-level* information, while religious concepts are mostly based on violations of *domain-level* information, which may engage different processes and lead to different recall effects.

To sum up, studies of memory effects on cultural material still leave open three questions. One is whether domain-level semantic incongruities of the type generally found in religious concepts trigger distinctive recall, which would help us understand why they are culturally widespread. A second question is whether these effects can over-ride culturally familiar, official theologies, as Barrett & Keil's results suggested. A third question is whether the recall effects of domain-level violations are found in different cultures, as suggested by the anthropological model described here.

To address these questions we used a series of free-recall tasks with essentially similar material in three different cultural settings, in France, Gabon and Nepal. We focused on the two categories that are most frequent in religious concepts, that of 'person' and 'artifact'. Although religious systems include many other types of concepts (mountains, rivers, plants or clouds with supernatural properties) these two are by far the most frequent (Boyer 1996). The stimuli consisted of quasi-stories, adapted from Barrett's studies of serial transmission (Barrett, 1996), in which a narrative frame brackets a list of descriptions of various intuitive and counterintuitive situations (see Appendixes 1 and 2). Such narratives have a beginning and an end, there is a main character and something happens to him, thereby meeting minimal intuitive criteria for a "story" (Mandler, Scribner, Cole & DeForest, 1980; Brewer & Lichtenstein, 1981; Ackerman et al., 1990). The format allowed us to produce a list of different situations none of which has any particular effect, disruptive or otherwise, on narrative structure. This quasi-story was suitably modified for other cultural settings, allow-

ing us to use a task that made sense to non-Western participants.[2] We only used items that triggered no direct association with culturally familiar religious concepts.

2. Experiment 1

The first experiment contrasted standard situations with *breaches* of intuitive domain-level expectations: (i) nonstandard physical properties and (ii) nonstandard psychological properties. To evaluate the respective contributions of oddity and cultural familiarity, we also used two questionnaires to elicit (i) judgments of "oddity" for the situations described and (ii) judgments of familiarity, that is, whether participants thought they had previously encountered these situations in stories, films, cartoons, and so forth We predicted a strong correlation between recall performance and the results of the first questionnaire (the distinctiveness literature suggests that items that are perceived to be odd are also recalled well) but no correlation with the second one (we conjectured that this effect would not be affected by how familiar the items were).

2.1. Method

2.1.1. Participants

Participants were recruited from humanities undergraduates (in various disciplines excluding psychology) at Université Lumière, Lyon. They took part in this experiment as the first part of a paid one-hour session. There were 18 participants (12 women, 6 men), aged 18 to 33 ($M = 22$, $SD = 3$).

2.1.2. Materials and norming

The recall material consisted in a two-page story adapted from Barrett (1996). A diplomat is about to be sent as an ambassador to a distant galaxy. He goes to the local museum to get a better idea of what to expect over there. Between introduction (arrival at the museum) and end (return home), the main part of the text is a list of 24 short descriptions of exhibits in the museum. We used four versions of the stories, with identical items in different orders.

The story included 12 items for each of the Artifact and Person categories. In each category, six items described a breach of intuitive expectations ("Br" items) and the other six an expected association ("Sn" items). Person items consisted in colloquial descriptions of *psychological* features taken from the "theory of mind" literature, for example, "there were people who could see what was in front of them" (Sn) or "there were people who could see through a wall" (Br). The artifact items described *physical* features taken from the "intuitive physics" literature, for example, "pieces of furniture that you can move by pushing them" (Sn) and "pieces of furniture that float in the air if you drop them" (Br).[3] Order of presentation was counterbalanced.

All items were pretested with 18 students and staff of Université Lumière in three different conditions: asking them to rate the items as "normal v. abnormal," "banal v. surprising" and "familiar v. unfamiliar." Items that reached less than 90% consensus were discarded. The remaining items were slightly modified to result in a similar sentence structure and roughly

similar word-count. In the story, all items were described in two sentences, the second one being a straightforward paraphrase of the first.

A first questionnaire form (Questionnaire 1–1) included all the items in the recall text, in a different order. Instructions read as follows: "The following sentences describe the different exhibits Mr. Wurg saw in the museum. For each sentence, indicate whether you find the object or person described similar or different from what we usually encounter in reality, by checking the appropriate box."

A second questionnaire (1–2) had the same list of items and different instructions: "The following sentences describe the different exhibits Mr. Wurg saw in the museum. Indicate whether each sentence describes an object or person that you have previously encountered, either in reality or in films, stories, cartoons, and so forth by checking the appropriate box."

2.1.3. Design and procedure

This was a 2 (category) x 2 (level) design with both category (artifact vs. person) and level (standard items vs. breaches of expectations) as within-subject variables. The participants were all tested individually in an experimental booth. They were given the printed text of the story. The instructions were to read the story very carefully and try to imagine each situation described in the story. When they reached the end of the text they were instructed to hand back the text to the experimenter. As a distraction task, they were then asked to do some mental arithmetic and to multiply the number of vowels in various words. After this 5 min distraction task, they were then given ruled sheets and instructed to write down as many of the exhibits in the museum as they could recall without regard for item order but with as much detail as they could recall. This part of the test was limited to ten minutes.

They were then given the two questionnaires and instructed to give spontaneous, literal responses and avoid metaphorical interpretations of the items. For the second questionnaire (1–2), the experimenter first asked the participant to explain the difference with questionnaire 1–1 and used two "training" questions to check that both fictional and real familiar items were assigned to the same category. The participants were then de-briefed about the aims of the experiment and asked whether they saw any connection between the items and religious notions.

2.2. Results

2.2.1. Recall

The results are summarized in Fig. 1. Recall scores were raw number of items recalled by each participant in each cell. An item counted as recalled if the participants' version included (i) what made it a member of the stimulus categories and (ii) what made it different from other items.[4] Overall mean recall was 42.59% or 10.22 items out of 24 ($SD = 3.56$). A 2 (level) x 2 (category) ANOVA showed a significant effect of level (standard descriptions vs. breaches of expectations), F (1, 17) = 26.5, $p < .0001$, with better recall for Br items, a significant effect of category, F (1, 17) = 22.22, $p < .0001$, with better recall for Artifact items than Person items. There was no significant interaction between category and level, F (1, 17) = 2.07, $p = .167$.

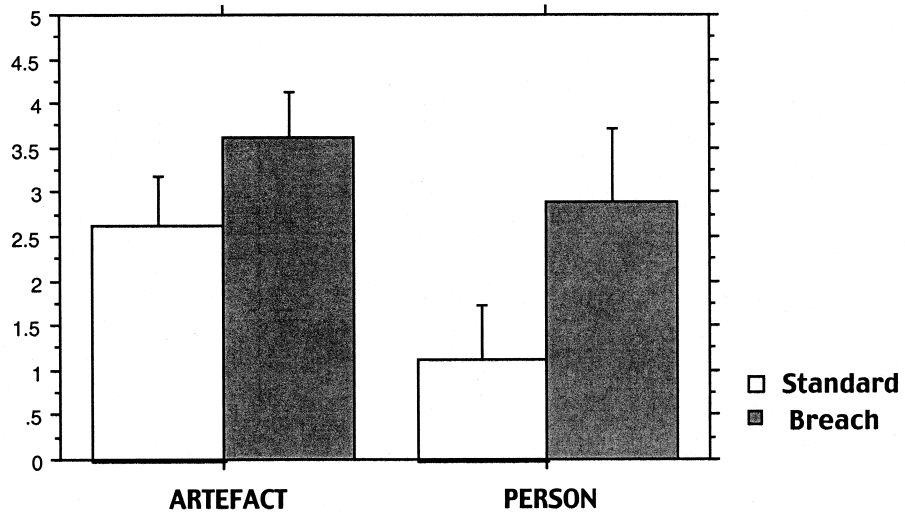

Fig. 1. Experiment 1. Mean number of Standard and Breach items recalled (max. 6) in the Artifact and Person categories, with 95% confidence intervals.

2.2.2. Questionnaire results

They are summarized in Table 1. For questionnaire 1–1 (whether items are "different from what we find on Earth") a 2 (level) x 2 (category) ANOVA showed a significant effect of level, F (1, 68) = 546, $p < .0001$, a significant effect of category, F (1, 68) = 8.05, $p = .006$, as well as a significant interaction between level and category, F (1, 68) = 4.6, $p = .036$. For questionnaire 1–2 (whether items were "ever encountered before, either in fiction or in reality"), a 2 (level) x 2 (category) ANOVA showed a significant effect of level, F (1, 68) = 128, $p < .0001$, no effect of category, F (1, 68) < 1, and no interaction between category and level, F (1, 68) < 1.

2.2.3. Correlation

There was a significant correlation between recall performance and the results of questionnaire 1–1 (items judged "different"), $r = 0.474$, $p < .0001$, and a significant correlation between recall and items judged "unfamiliar," $r = 0.373$, $p = .001$.

Table 1
Experiment 1, Questionnaires 1-1 and 1-2. Mean number (and *SD*) of items judged "different from what is on Earth" and judged "not encountered before in fiction or reality" for Sn items (no breaches) and Br items (breaches) in two categories, Artifact and Person.

		Standard	Breach
judged "different"	ARTIFACT	.28 (.826)	5.67 (.97)
	PERSON	.28 (.461)	4.5 (1.09)
judged "unfamiliar"	ARTIFACT	0 (-)	3 (1.75)
	PERSON	.06 (.24)	2.72 (1.18)

2.3. Discussion

Recall results support the hypothesis that sentences that include a breach of expectations are recalled better than standard ones. Recall results also showed a category effect. Artifact items were recalled better than Person items in both Sn and Br item-types. This could be caused, either by an intrinsic advantage of artifact sentences or by an advantage of physical properties. In this design all the properties of artifacts were *physical* properties and all those associated with persons were *psychological* properties. Properties in a study of *breaches* cannot be fully crossed between categories, because artifacts do not include psychological or biological properties by default. Whether the effect was due to the properties or to the categories themselves is addressed in Experiment 2.

The results of Questionnaire 1–1 also showed both level and category were significant factors in judging that items are "different" from what is found on Earth. The level effect was expected, since the stimuli had been assigned to the Br or Sn cells on the basis of very similar pretest questionnaires. The category effect showed that participants' judgments about the strangeness of breaches were less definite for psychological properties than for physical ones. This may help towards explaining the category effect in recall. If the participants found it easy to imagine counterintuitive psychological properties, these Br items would be less distinctive compared to Sn items than physical Br properties compared to Sn ones. There was an important divergence between the results of the two questionnaires, in the predicted direction. These results suggest a strong correlation between the fact that items are explicitly rated as "different" from what is found on Earth and the likelihood that they are remembered in the free recall task.

None of the participants saw any connection with religious concepts before the experimenter de-briefed them.

Overall, these results support a "distinctiveness" interpretation of recall for semantically incongruous material. However, they also raise the question of what makes the Br items distinctive. It could be that the combinations of categories + properties were incongruous (our hypothesis) or that the properties by themselves were unfamiliar. Both interpretations are consistent with these results. This question is addressed in Experiment 2, where all properties are familiar ones.

3. Experiment 2

The point of Experiment 2 was to test the hypothesis that violations are recalled better than standard associations even if the properties themselves are familiar. This occurs in a *transfer* of properties across categories, a phenomenon that is common in religious representations (Boyer, 1996b). An object or being is described as belonging to a particular category yet has properties usually excluded by membership in that category, for example, an artifact with cognitive capacities (e.g., a statue that listens to people) or a person with artifact properties (golems and other person-like beings assembled by someone). Such transfers do not require that one represents unfamiliar or impossible properties. In this way, Exp. 2 could test whether the results of Exp. 1 were caused by the incongruity of the situations described or by the

oddity of the properties themselves. This experiment used similar properties for both artifact and person items, which should make it possible to evaluate the category differences observed in Exp. 1. The main hypothesis was that transfer items would be recalled better than no-transfer items. The story used was similar to that of Exp. 1, as were the two questionnaires following the recall test.

3.1. Method

3.1.1. Participants

Twenty-two Humanities, nonpsychologist majors (12 women, 10 men, aged 18 to 34, $M = 21.6$, $SD = 3.73$) at Université Lumière, Lyon were recruited as paid participants as part of a one-hour session.

3.1.2. Materials

The story frame was similar to that used in Exp. 1. The stories included items from two categories, Artifact and Person. For each category, there were six "Tr" items describing a *transfer* of expectations, that is, a category with a description of a feature that is usually expected for another category. Six "Sn" items described properties that could be expected in members of that category. All sentences had been normed in the same way as in Experiment 1. There were four different versions of the stories with identical items in a different order.

For the Person category, Sn items described normal psychological properties. Tr items described the way the person had been made, for example, "people who are made of a rare metal," "people you can put together only with special tools," and so forth For the Artifact category, psychological properties were Tr items and "manufacturing" properties were the Sn ones.

3.1.3. Design and procedure

This was a 2 (category) x 2 (level) design with both category (artifact vs. person) and level (standard items vs. transfers of expectations) as within-subject variables. The procedure, as well as instructions for both questionnaires, were similar to those of Experiment 1.

3.2. Results

3.2.1. Recall

Mean recall rate was 49.43% or 11.86 items out of 24 ($SD = 3.52$). Results of the recall task were scored in the same way as for Exp. 1. Recall data are summarized in Fig. 2 below. A 2 (level) $\times$ 2 (category) ANOVA showed a significant effect of level (standard items vs. transfers of expectations), F (1, 21) $= 4.84$, $p = .039$, with higher recall rate for transfer items, a trend for category (person vs. artifact) below significance, F (1, 21) $= 3.264$, $p = .085$, and no significant interaction between level and category, F (1, 21) < 1.

3.2.2. Questionnaire results

They are summarized in Table 2 below. For 2–1, whether items were "different" or not from what is found on Earth, a 2 (level) $\times$ 2 (category) ANOVA showed a significant effect

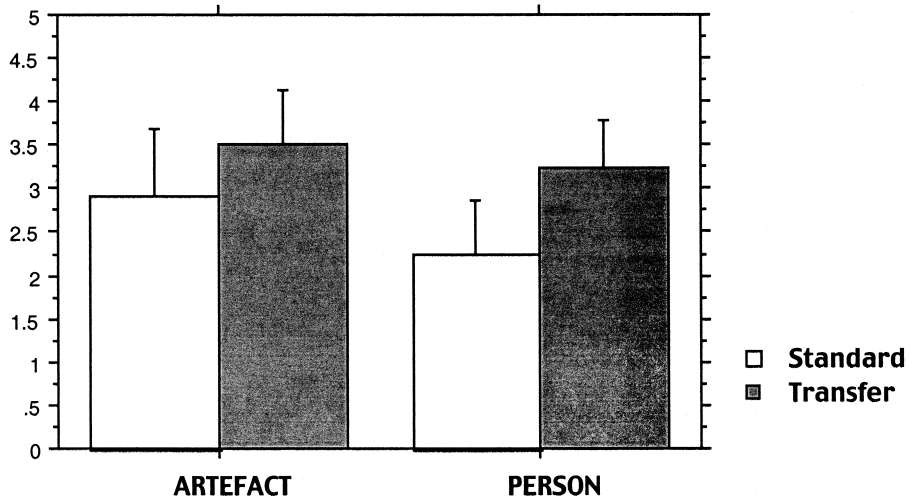

Fig. 2. Experiment 2. Mean number of standard and Transfer items recalled (max. 6) in the Artifact and Person categories, with 95% confidence intervals.

of level, F (1, 84) = 359, $p = < 0.0001$, no significant effect of category, F (1, 84) < 1, and no interaction between category and level, F (1, 84)<1. For questionnaire 2–2, where items were judged as "encountered before in fiction or reality" or not, a 2 (level) x 2 (category) ANOVA showed a significant effect of level, F (1, 84) = 39.8, $p = <0.0001$, a trend but no significant effect for category, F (1, 84) = 3.47, p = .066, and a significant interaction between level and category, F (1, 84) = 10.88, p = .0014.

3.2.3. Correlation.

There was a significant correlation between recall performance and the results of questionnaire 2–1 (items judged "different," r = 0.255, p = .016, but no significant correlation between recall and questionnaire 2–2 (items judged "unfamiliar"), r = 0.035, p = .75.

3.3. Discussion

Recall results supported the hypothesis that items including a transfer of predicates from another category would be better recalled than items without such transfers. As expected, the

Table 2.
Experiment 2, Questionnaires 2-1 and 2-2. Mean number (and *SD*) of items judged "differenct from what is on Earth" and judged "not encountered before in fiction or reality" for Sn items (no transfers) and Tr items (Transfers) in two categories, Artifact and Person.

		Standard	Transfer
judged "different"	ARTIFACT	.273 (.703)	5.14 (1.75)
	PERSON	.045 (.213)	5.05 (1.53)
judged "not encountered"	ARTIFACT	.455 (1.08)	1.409 (1.65)
	PERSON	0 (-)	3.05 (2.24)

results of Questionnaire 2–1 correlate strongly with those of the recall task. The results of Questionnaires 2–1 and 2–2 show that many items are judged both "nonstandard" and "encountered before," suggesting that recall is associated with distance from intuitive expectations more than with cultural familiarity.

Overall, Experiments 1 and 2 suggest that violating expectations or transferring them across categories produces a "distinctiveness effect" in a free recall task. In both studies we find that violations were recalled significantly better than standard items. This would be in keeping with the contention that "bizarreness" or "distinctiveness" effects are an effect of *expectation-violation* (Hirshman, 1988). The items are compared with conceptual schemata and the discrepancy is sufficient to produce the effect. Waddill & McDaniel speculate that a comparison between target sentences and prior knowledge could occur very quickly during processing and tag atypical items with episodic information that boosts retrieval (Waddill & McDaniel, 1998, p. 118). Such effects would support the anthropological conjecture, that recall effects may contribute to the cultural spread of violations of information at the domain level.

4. Experiment 3

Our results so far show that recall favors both breaches (properties that are appropriate for the category but violate expectations) and transfers (properties that are not appropriate for the category). This might suggest that combinations of these factors would trigger even stronger recall effects by presenting people with more obviously "supernatural" concepts. However, religious concepts rarely include such complex violations. In particular, combinations of breaches and transfers are virtually nonexistent in religious concepts, even in cultural settings where both breaches and transfers are common. For instance, many Catholics in Europe have representations of statues that listen (transfer), as well as a representation of a God who can hear sounds from anywhere in the world (breach). But they do not assume that listening statues can perceive distant sounds; people who pray to such statues generally stay within hearing distance, as it were, of the artifact.

That both breaches and transfers are common but not their combination cannot be a simple matter of conceptual overload, as the Christian God, for instance, combines many breaches (prescience, ubiquity, eternity, etc.). Our anthropological model would suggest that combinations are rare because they block inferences usually provided by the ontological category (in common conceptual associations) and preserved in the case of limited violations (breaches and transfers). If so, one would *not* expect the breach-transfer combinations to be recalled better than either breaches or transfers, although they are more distant from common associations.

Experiment 3 was designed to test whether recall also contributes to this other important feature of religious concepts. To do this, we tested free recall for items that combine the two factors studied so far. That is, they use properties that are intrinsically strange in their domain of application and apply it to an inappropriate domain. To allow comparisons with Experiments 1–2, we used essentially similar material and combined it to produce the following cells:

1. standard, no breach, no transfer ("Sn" items),
2. breach with no transfer ("Br" items),
3. transfer with no breach ("Tr" items),
4. breach with transfer ("BrTr" items)

A BrTr item includes a counterintuitive predicate *and* applies it to a category for which it is inappropriate. Consider for instance "only remembering what did not happen." This is a counterintuitive psychological property for any being or object that has a psychology: a person, but also an animal or an intelligent computer. If this kind of breach predicate is applied to categories of objects which do not normally have psychological processes (e.g., a piece of furniture), a transfer is added to the breach. Since we had some evidence from Experiments 1–2 that both transfer and breach contributed to better recall, this study could indicate how their effects are enhanced or cancelled when these factors are combined.

4.1. Method

4.1.1. Participants

Twenty-one undergraduate students at Université Lumière, Lyon (12 women and 9 men, aged 18 to 25,$M = 21$, $SD = 2.1$) took part as part of a paid hour-one session.

4.1.2. Materials

The narrative frame was the same as in Exp. 1 and 2. All the "exhibits" in the museum described properties of artifacts: standard physical properties (Sn), breaches of physical expectations (Br), standard psychological properties (Tr), breaches of psychological properties (BrTr). Some items were rephrased so that BrTr items were neither longer nor syntactically different from other item-types. We used four different versions to vary order of presentation with identical items.

4.1.3. Design and procedure

The main design was a 2 (breach) x 2 (transfer) design in which both category and level variables were manipulated within subjects. The procedure for recall tests as well as the questionnaire instructions were identical to those of Exp. 1-b.

4.2. Results

4.2.1. Recall

Overall mean recall was 40.01% or 9.62 items out of 24 ($SD = 2.96$). Fig. 3 breaks down recall rates by item-type.

A 2 (breach vs. no-breach) x 2 (transfer vs. no-transfer) ANOVA showed a significant effect of breach, $F(1, 20) = 4.47$, $p = .0471$, with higher recall for Breach items, no overall effect of transfer, $F(1, 20) < 1$, and a significant interaction between breach and transfer, $F(1, 20) = 10.28$, $p = .0044$. Because of this interaction, we carried out a series of planned comparisons, summarized in Table 3.

These comparisons confirm that whether the property itself is counterintuitive (breach vs.

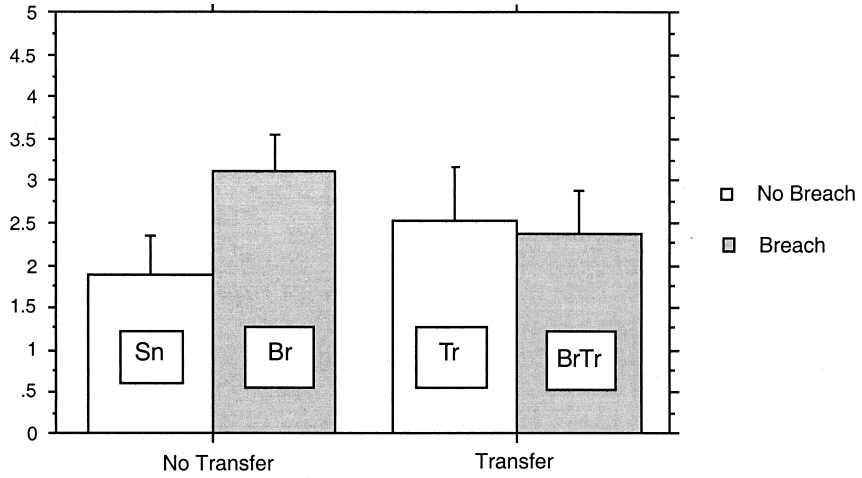

Fig. 3. Experiment 3. Mean number of No-Breach and Breach items recalled (max. 6), both No-Transfer and Transfer, with 95% confidence intervals.

no breach) has an effect only when the properties are appropriate for the category (no transfer). Whether the property is appropriate (Transfer × No-Transfers) has an effect on both intuitive and counterintuitive properties, but it goes in opposite directions. Transfer of a No-Breach property increases recall while transfer of Breach properties decreases recall.

4.2.2. Questionnaire results

They are summarized in Table 4. For 3–1 (whether items were judged "different from what we find on Earth"), a 2 (breach vs. no-breach) 2 (transfer × no-transfer) ANOVA showed a significant effect of breach, F (1, 80) = 80, $p < .0001$, a significant effect of transfer, F (1, 80) = 152, $p < .0001$, and a significant interaction, F (1, 80) = 76, $p < .0001$. For questionnaire 3–2, (whether items were judged "encountered before, either in fiction or reality"), a 2 (breach vs. no-breach) × 2 (transfer vs. no-transfer) ANOVA showed a significant effect of breach, F (1, 80) = 51.41, $p < .0001$, a significant effect of transfer, F (1, 80) = 41.5, $p < .0001$, and no significant interaction, F (1, 80) < 1.

4.2.3. Correlation

There was a significant correlation between recall performance and the results of questionnaire 2–1 (items judged "different"), $r = 0.341$, $p = .0014$ but no correlation between recall and questionnaire 2–2 (items judged "unfamiliar"), $r = 0.057$, $p = .61$.

Table 3
Experiment 3. Summary of planned comparisons.

	MS	*F*(1, 20)	*p*
1. No Transfer: Breach vs No Breach	14.88	16.38	.0006
2. With Transfer: Breach vs No Breach	.214	.236	.633
3. No Breach Items: Transfer vs No-transfer	4.02	4.43	.0482
4. Breach items: Transfer vs No-Transfer	5.37	5.898	.0247

Table 4
Experiment 3, Questionnaires 3-1 and 3-2. Mean number (and *SD*) of items judged "differenct from what is on Earth" and judged "not encountered before in fiction or reality" for four categories of items.

		No breach	With breach
judged "different"	No transfer	.048 (.218) [Sn]	4 (1.095) [Br]
	With transfer	4.76 (1.09) [Tr]	4.81 (1.33) [BrTr]
judged "not encountered"	No transfer	.143 (.478) [Sn]	1.952 (1.396) [Br]
	With transfer	1.762 (1.411) [Tr]	3.71 (1.271) [BrTr]

4.3. Discussion

The recall results showed an interaction between the two variables crossed to produce four categories of stimuli. As a result, properties that are intrinsically counterintuitive are recalled better if they are appropriate for the category, less if they are not. In contrast to possible extrapolations from the results of Experiments 1 and 2, combinations that represent the most salient departure from expected associations are not recalled better than simple breaches or transfers.

Decreased recall for material that is further removed from common schemata could be caused by the fact that BrTr items were (i) so strange they were "normalized" to Sn items by the subjects or (ii) misconstrued as Br or Tr items expressed in a complicated way or (iii) less clearly represented because of their inherent complexity. These explanations would predict confused renditions of BrTr items, which was not the case. In this as in Exp. 2, we found so few distorted items (or transfers to another cell, e.g., Sn items turned into Br) that their analysis was not possible. Besides, BrTr items were generally judged "different" from common representations and the most "unfamiliar" of all item-types. Compared to the other experiments, we find here a strong divergence between explicit judgment and recall. While BrTr items are rated "different" and "unfamiliar" more than any other item-types they are recalled less than either simple breaches or simple transfers.

That recall is optimal for limited violations and decreases with more bizarre material is a familiar result in memory research, and more generally illustrates Kagan's "stimulus-schema discrepancy" model, following which attention and recall are enhanced by limited departures from activated schemas but decrease as items are distorted beyond a certain limit (see e.g., Zelazo & Shultz, 1989; Bloom, 1998). This effect was demonstrated in a variety of attentional and memory tasks for verbal and visual material. In this view, more bizarre material may result in decreased recall because it makes it less likely to activate the relevant schema. However, this operational understanding of "discrepancy" may denote very different phenomena, depending on the kind of "schema" activated. In the present case, what makes Breach+Transfer items different from other types is not that the relevant domain-level expectations were not activated, but that inferences on the basis of these expectations were blocked. An artifact that goes right through walls (Breach) still preserves some intuitive features of artifacts (it was made by people, for some purpose, etc.) and of physical objects (it has a location in space, a continuous trajectory, etc.). An artifact that thinks or has feelings (Transfer) still maintains intuitive features of solid objects and allows further inferences from psychological expectations (it has perception, memory, etc.). This latter domain of inferences

is blocked in the case of BrTr items, described as having nonstandard psychological properties. One may speculate that this is one of the factors that lead to decreased recall as well as cultural rarity of such combinations of breaches and transfers.

Taken together, the results of experiments 1–3 would seem to suggest that cultural familiarity is not a major factor in recall for limited violations of domain-level information. However, this last hypothesis was only supported by explicit judgments and these originated from a single cultural setting. To test this divergence between recall and cultural familiarity, one should test such material in different settings.

5. Experiment 4

To test the cross-cultural validity of the conclusions from Experiments 1–3, we chose to conduct a series of similar studies in a cultural setting where one of us had conducted anthropological fieldwork (Boyer, 1990). This was a different context in terms not just of social and economic conditions but also of (i) people's everyday familiarity with religious concepts, (ii) the range of such concepts and (iii) their mode of transmission. By contrast to the secularized milieu in which our French participants live, most inhabitants of Libreville, Gabon, are familiar with religious concepts and occurrences. That is, many events, either trivial or consequential, are readily construed as caused by supernatural agents. Second, concepts of such agents are much more diverse than in a Western context. Concepts from Christian denominations are completed with local witchcraft concepts and elaborate notions of ghosts and shadows (Boyer, 1990). Third, information concerning such agents is acquired in the context of informal social interaction, not through literate sources or communication with specialized scholars.

Although the items were chosen from the same lists as in Exp. 1–3, this was not conducted as a straightforward replication, because of the special circumstances of this study. First, for this cultural setting we did not have a wealth of experimental results showing a recall effect for basic-level violations, as we did for Western subjects (see references to "distinctiveness" in the Introduction). We could not simply assume that such effects obtained with our Gabon participants, since one of our aims was precisely to test the cross-cultural validity of the schemata that make certain representations distinctive. We remedied this by adding a level of "basic-kind-level" violations to our lists. Our stories included items describing standard situations ("Standard" items, e.g., "a man who was slightly taller than a woman"), violations of kind-level information ("Kind-incongruous" items, e.g., "a man who could uproot a tree with his bare hands"), and category-level transfers or breaches ("Domain-incongruous" items, e.g., "a man who could walk right through a mountain"). Second, because of the increased number of items, we chose to manipulate category as a between-subject variable. Third, it proved impossible to use questionnaires to test the participants' judgments of cultural familiarity for the items, because of the sensitive nature of the topic.[5] We remedied this by pretesting the items with assistants and eliminating all items that could evoke particular folk-tales or witchcraft themes.

5.1. Method

5.1.1. Participants

All participants were recruited informally at Libreville farmers' markets by a team of research assistants. Conditions for selection were (a) age (between 16 and 35), (b) French as first language and (c) educational level. We excluded participants who were either professionally engaged in literate work or deeply involved in religious activities. Participants were offered a small fee (a soft drink) for their participation. There were 81 participants (39 women, 42 men), age ranging from 16 to 37 ($M = 24.3$, $SD = 4.59$).

5.1.2. Materials

The stories were similar to those used in Exp. 1-b, with the exception of the narrative frame that bracketed the list of items. In the stories used here, an orphan leaves his village to seek fortune and returns empty-handed but with an account of all the villages he has visited. Stories of this kind are very common in Central Africa. There were two versions of the story, with Artifact and Person items respectively, each including six "Standard," six "Kind-incongruous" and six "Domain-Incongruous" items.

The items were taken from the lists used with the Lyon participants. They were pretested with a group of 10 undergraduate students at Libreville University, in the same way as in Exp. 1-b. We also asked subjects whether they could think of religious rituals or magical activities or folk-tales that included persons or artifacts described in the list and removed items that triggered such associations.

5.1.3 Design and procedure

This was a mixed 2 (category) x 3 (level) design, with category (Artifact vs. Person) a between-subject variable and level (Standard, Kind-incongruous, Domain-incongruous) a within-subject variable. Pilots showed that it was not desirable to conduct the experiment in the special rooms available on campus. In this culture being taken to a secluded place would be highly unnatural and suspicious. So a team of linguistics and anthropology students at the University of Libreville visited various markets and chose a quiet place where they could work in people's sight but without too much interference. Our assistants had been trained in interview techniques and the identification of first-language.[6] They tested the participants one by one, asking them first various questions about occupation, educational level, age and first language. They then read the stories, presented as a test for text-comprehension. (Pilots showed that people explicitly tested for memory would generally refuse to take part). After reading the stories, the experimenters asked them to perform elementary sums. They were then asked to relate what happened in the story in as much detail as possible. This was recorded by the experimenter. The subjects were then all de-briefed about the purpose of the experiment.

5.2. Results

Recall was scored in the same "conservative" way as in Exp. 1-c and 2. We discarded confused renditions and distortions that included material from several items. Overall mean

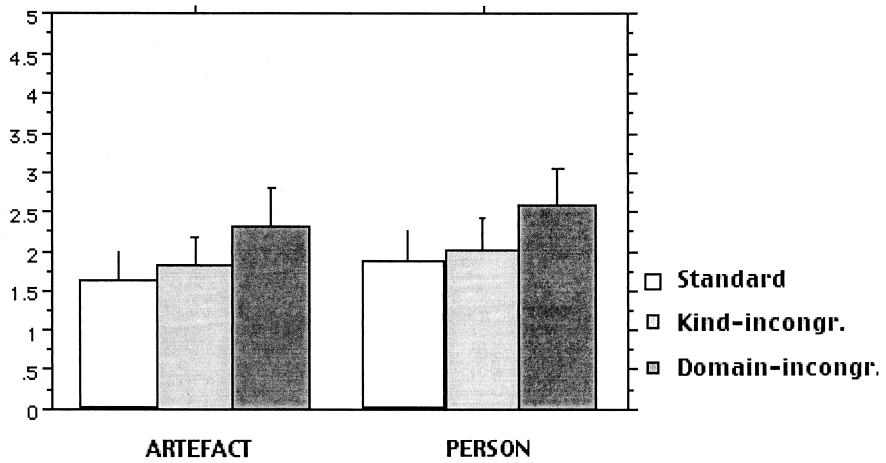

Fig. 4. Experiment 4. Mean number of items recalled, in two conditions (Artifact and Person) for three categories of items: Standard, Kind-incongruous and Domain-incongruous.

recall over both conditions was 31.62% or 6.075 items out of 18 ($SD = 2.58$). Mean recall was 6.43 items ($SD = 2.86$) in the Person condition and 5.69 items ($SD = 2.2$) in the Artifact condition.

A mixed 2 (category) x 3 (level) ANOVA on these results showed no effect of category, F (1, 79) = 1.66, $p = .201$, and a significant effect of level, F (2, 79) = 8.39, $p = .0003$. There was no interaction between category and level. However, there were some very low scores (one or two items recalled) that suggested that some participants had not grasped the point of the experiment, which was confirmed by subsequent interviews. To check whether these would produce spurious effects, we performed an additional ANOVA on recall performance exclusive of participants with low scores (participant's total number of items recalled $<(M - 1\ SD)$ in each condition). This produced similar results: no effect of category, F (1, 70)$<$0.41, a significant effect of level, F (2, 70) = 9.4, $p = .0001$, and no significant interaction between level and category, F (2, 70) = 0.23.

Fig. 4 illustrates these results, excluding very low scores.

We also ran a series of related t tests to produce pairwise comparisons between conditions, summarized in Table 5.

5.3. Discussion

These results support the main hypothesis, that violations of domain-level information trigger high recall for both artifact and person concepts. The main result was the significant difference between "Standard" and "Domain-incongruity" items, supporting the hypothesis that violations at the domain level trigger better recall than standard conceptual associations, in this cultural setting in the same way as with our French subjects.

Taken together, the results of Exp. 1–4 suggest that cultural differences have no noticeable effect on recall for category-level violations. Experiment 4 showed that neither everyday familiarity with concepts that include violations, nor a greater variety of such concepts,

Table 5
Experiment 4. Summary of comparisons between levels of recall for different item-types in the Artifact and Person conditions.

		$t(38)$	p
	Standard vs. Kind-incongruity	.407	.6865
ARTIFACT	Kind- vs. Domain-incongruity	2.399	.0225
	Standard vs. Domain-incongruity	2.639	.0116
		$t(32)$	p
	Standard vs. Kind-incongruity	.884	.382
PERSON	Kind- vs, Domain-incongruity	2.44	.019
	Standard vs. Domain-incongruity	2.588	.0136

influence such effects, once the information that makes these violations familiar is removed. Finding similar recall effects across cultures suggests, first, that the links between the categories and the predicates used are represented in a similar fashion in these two cultures. This in itself is not too surprising, in the sense that pretesting of the items revealed very few differences in judgment between French and Gabonese participants. We did not expect the person or artifact categories to be so different that such properties as “can see through walls” (for person) or “is sad when it is left alone” (for artifact) would be judged “normal” or “usual” in Gabon. More important, these results also show that such similar links between categories and properties result in similar recall performance for violations of those properties.

However, there might be a bias here. Despite cultural differences, our French and Gabonese groups were both taken from populations with minimal exposure to scholarly, literate religion. These were representative of both populations, though for different reasons (secularization in France, mainly oral transmission in Gabon). Now it could be argued that the cross-culturally similar sensitivity to violations observed here might be modified by the presence and cultural salience of scholarly religion, a question addressed in Experiment 5.

6. Experiment 5

This was conducted in a cultural setting where one of us had conducted anthropological fieldwork (Ramble, 1984). The population from which we selected participants (Tibetan monks in Katmandu) offered several advantages for this study. First, this is a place where religious concepts are generally influenced by the canonical teachings of specialized Lamaist institutions. Second, the monks we worked with are experts in such concepts. They spend much of their time studying sacred texts that give explicit descriptions of the special characteristics of religious agents and situations. Their tradition (*Bon*) is broadly similar to Tibetan Buddhist tradition[7] in that it places great emphasis on ways to change one’s cognitive and emotive processes in order to reach a higher form of consciousness. A whole range of exercises, for example, meditation, sensory deprivation, altered states of consciousness, and so forth, are used to effect such transformations (Ramble, 1984). The texts monks

read and study include many accounts of extraordinary feats performed as a result of such mental training.

This offered us an opportunity to test the relative validity of two interpretations of the cognitive effects of such literate theological versions of religious concepts. Most anthropological models, while noting the differences between official and popular theologies, assume that literacy has a profound influence on both "official" descriptions of religious concepts and their "popular" form (Goody, 1977; Goody, 1986). By contrast, Barrett and Keil have presented suggestive evidence that the effects of official theologies are limited when task-demands require fast, "on-line" inferences about religious agency (Barrett & Keil, 1996; Barrett, 1998). Our experiment was relevant to another aspect of this question, that of the relative salience of *types* of violations used in a theology. Since person violations are a privileged theme in this cultural setting, one could expect that items including such violations would be better recalled than artifact items, if familiarity with theological themes was an important factor driving recall. If, on the other hand, violations of intuitive expectations were the main factor, then artifact violations would be advantaged by the fact that they belong to less familiar templates.

6.1. Method

6.1.1. Participants

All participants were monks of the Triten Norbutse monastery in Katmandu. They were all native speakers of Tibetan. Participants were selected on grounds of age and literacy. There were 30 participants, all male, age ranging from 14 to 30 ($M = 22.45$, $SD = 4.6$). Although this spans a wide age-range, all participants were more or less at the same stage in their curriculum, that is, they were all novices or young monks with roughly equivalent experience of monastery teaching and the consequent familiarity with literacy (at least four years).

6.1.2. Materials

The recall materials, as in Exp. 1-c, consisted in a quasi-story with an embedded list of 18 two-sentence descriptions of different situations. The narrative frame from Exp. 1-c was modified to use a culturally familiar format. In this story, a trader returns to his village and tells his friends of all the things and people he has encountered. There were two versions of the story, one with descriptions of people with particular characteristics (PER items) and the other one with artifacts (ART items) in all three levels. Item order was then counterbalanced to produce four different versions of each story.

Both stories were translated into Tibetan by C. Ramble with the help of Tenpa Yungdrung, vice-abbot of the Triten Norbutse monastery, to achieve a sufficiently idiomatic Tibetan rendition of the items. Items were modified (i) if they were made confusing by local idioms, (ii) if they happened to correspond too closely to some local story or belief, or (iii) if they were just unintelligible given the cultural context. The stories were back-translated to check that the items were still essentially similar to those of Experiment 4. They were printed in Tibetan script and Xeroxed to produce handouts for the participants.

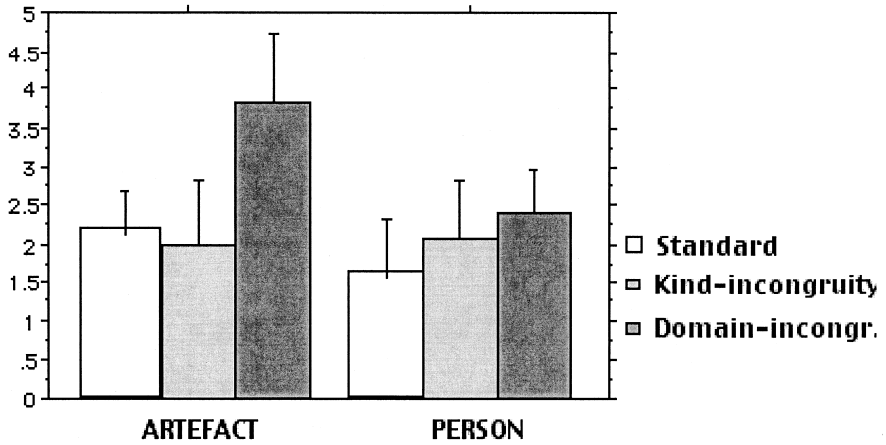

Fig. 5. Experiment 5. Mean number of items recalled, in two conditions (Artifact and Person) for three categories of items: Standard, Kind-incongruous and Domain-incongruous.

6.1.3. Design and procedure

This mixed design was similar to that of Experiment 4, with category (Artifact vs. Person) as a between-subject variable and level (Standard, Kind-incongruous, Domain-incongruous) as a within-subject variable.

The participants were gathered by the vice-abbot and informed that they would take part in an informal "test." They were not told that this would be a memory test. They were given the handouts and instructed to read them attentively. After about 15 min, they were asked to give the scripts back and to write down their name and age on a sheet of paper. As a distractor task, they were then asked to do sums and take down the results. They were then asked to write down all they could remember of the situations encountered by the hero in as much detail as possible, regardless of the order of items in the story. The experimenters then de-briefed the participants about the purpose of the experiment.

6.2. Results

An item scored as positive if participant's version included the particular details that made it a member of the category and distinguished it from other items in the story. This excluded a number of cases for which the participants had mixed information from several items. Overall mean recall over both conditions was 38.52% or 6.93 items out of 18 ($SD = 1.55$). The mean number of items recalled was 6.06 items ($SD = 2.9$) in the Person condition and 8.08 items ($SD = 2.69$) in the Artifact condition. Recall data are summarized in Fig. 5 below, with results broken down by category and level.

A mixed 2 (category) x 3 (level) ANOVA showed a trend for category with a slightly higher recall rate for Artifact items, $F(1, 29) = 3.79$, $p = .061$. There was a significant effect of level, $F(2, 29) = 5.13$, $p = .012$, with Ontological items best recalled, followed by Basic and Standard items in both categories. There was no significant interaction between category and level.

Table 6
Experiment 5. Summary of comparisons between levels of recall for different item-types in the Artifact and Person conditions.

		$t(13)$	p
	Standard vs. Kind incongruous	.144	.887
ARTIFACT	Kind incong. vs. Domain-incongr.	3.33	.0054
	Standard vs. Domain-incongr.	3.87	.0019
	Standard vs. Kind incongruous	1.33	.205
PERSON	Kind incong. vs. Domain-incongr.	.467	.648
	Standard vs. Domain-incongr.	1.75	.102

As in the Libreville study and for similar reasons, we performed an additional ANOVA on recall performance exclusive of low scores ($x<(M - 1SD)$ in each condition) to check for spurious effects induced by very low scores. This shows no effect or trend for category, $F(1, 21)<1$, a significant effect of level, $F(2, 21) = 5.7$, $p = .0065$ and a trend interaction, $F(2, 21) = 2.79$, $p = .078$.

Fig. 5 summarizes these results exclusive of very low scores.

As in the Gabon experiment, we also ran a series of related *t* tests to produce pairwise comparisons between conditions summarized in Table 6.

6.3. Discussion

These results support the main hypothesis, that items including domain-level violations are generally recalled better than common conceptual associations, while violations of kind-level information do not produce such effects. However, finer-grained analysis shows that the contrast between levels of violation is only observed in the artifact category. Indeed, in that category Domain-level violations are recalled significantly better than Kind-violations. This did not result from a general advantage for all artifact cells or for all nonstandard artifact items. The category-difference was mainly driven by very high recall in the Domain-level violation artifact cell, in contrast to the Gabon results where there was little difference between the two catgories.

This difference may be explained by the special cultural context of this study. Although person violations are common in the culture and in the specialized texts monks are familiar with, the participants recalled artifact violations much better than all other types of items. This suggests that cultural familiarity with a certain *type* of violations (what we called a "template" in Introduction) does not result in better recall. One could speculate that violations of a type that is *not* used in the participants' culture are more salient because they seem more "strange" or "novel" or "exotic." Note that this effect concerns the familiarity of *templates*, not *concepts*. In these studies, we took care to remove items that triggered direct associations with familiar concepts.

More generally, this specific result may add to the evidence provided by Barrett & Keil concerning the resilience of intuitive, on-line concepts in situations of literate, theologically-informed religion (Barrett & Keil, 1996). Although our subjects were trained to pay special

attention to person-violations, this did not offset the novelty and intrinsic counterintuitive quality of artifact-violations.

Compared with the French and Gabonese studies, this experiment confirmed that overall recall performance was not much affected by the conditions of the test, as the rate of recall ($M = 6.93$ items out of 18) is quite similar to that of Gabon participants ($M = 6.075$). In comparison with the French results, this would tend to suggest that recall was affected more by the familiarity of the task (our French participants were used to academic tests) than by literacy in general or by the presence of literate religion (texts that describe violations) in the cultural environment.

7. General discussion

The purpose of present studies was to examine the contribution of recall to the spread of religious concepts in different cultural settings. We used a familiar and robust method–differential recall in a within-list design–to test hypotheses concerning the structure of religious concepts. In particular, we tested recall for material that includes violations of domain-level information, since such violations are a common feature of religious concepts.

The results supported the hypothesis that violations (either *breaches* of relevant expectations for a domain or *transfers* of expectations from one domain to another) produce a retrieval advantage (Experiments 1, 2). This might have suggested that stimuli that combine such breaches and transfers, thereby producing material that is further removed from common conceptual associations than simple breaches and transfers, would be recalled better. The opposite is the case, which could explain why such combinations are not culturally widespread (Experiment 3). The fact that some items were identified as culturally familiar was not significantly correlated with recall (Experiments 1, 2, 3). This might indicate that sensitivity to violations of category-level information is not influenced by cultural familiarity.

Indeed, recall results show that the tendency to recall such material is stable across cultures (Experiments 4, 5). Differential recall for domain-level violations remains despite cultural differences in people's commitment to actual religious concepts, in the range of concepts (limited sources vs. various sources), or in the prevalent mode of transmission (literate vs. oral). This would suggest that some recurrent features of religious concepts may be explained to some extent by a specific sensitivity to violations of ontological category information. This sensitivity does not seem to be substantially modified by the participants' cultural background, their familiarity with or commitment to religious beliefs, the kind of religious concepts used in their cultures or their mode of transmission. The results also support the anthropological conjecture, that the reasons why some ways of representing supernatural agency are recurrent between cultures and stable in a culture are to be found, not in the specifics of cultural *concepts* but in the structure of more abstract *templates.* Each culture may or may not make use of a particular template. For instance, we find many notions of artifacts with intentional properties in modern European cultures but few concepts of artifacts with biological properties (apart from the occasional Madonna shedding tears on Good Friday).

One could have expected such cultural differences to boost recall for novel concepts that

use familiar templates. However, we did not observe this effect. On the contrary, we found that Tibetan monks, who are familiar with many person-template concepts, recalled artifact items from templates that are not used in their culture (Experiment 5). Paradoxically, they were much less likely to recall person-domain violations. This suggests that cultural use of one kind of template does not hinder sensitivity to other ones, adding support to Barrett & Keil's studies of religious concepts (Barrett & Keil, 1996). The protocols were different, as Barrett & Keil studied distortions in recall (not straight recall) for religious agents (in contrast with our nonreligious items). However, the results converge in suggesting that a general sensitivity to certain types of supernatural concepts is only *partly* influenced by official and explicitly accessible theologies.

We did not observe differences related to the varying level of literacy (high in France and in this specific Tibetan population, low among our Gabonese participants). This is somewhat surprising given the well-established influence of literacy on many conceptual tasks, including recall. A possible interpretation would be that the present tasks are not really sensitive enough to reveal such differences (nor were they designed to test them). The massive effect of distinctiveness for domain-level violations might conceal some consequences of literacy. Also, the effects of literacy may be more important when participants are tested with the kind of material that they usually find in written sources (e.g., sacred texts for the Tibetan monks). (This was excluded here given our decision to use nonreligious material). So the precise effects of literacy might be illuminated by running further studies with material that is much closer to familiar sources.

In Experiments 4 and 5 we also found a step-wise recall effect with kind-level violations recalled better than standard associations but less than domain-level violations. These results are far from conclusive, but they suggest that violations of expectations may have different cognitive effects, depending on the level of conceptual information that is affected. Indeed, violations of domain-level information trigger specific intuitions of sentence anomaly (Gerard & Mandler, 1983) and specific event-related potential signatures (Polich, 1985) that do not occur with syntactic anomalies (Ainsworth-Darnell, Shulman & Boland, 1998).

Why are *limited* violations optimal? Recall for violations illustrates familiar effects of incongruity. Bizarre items are recalled better both because they are "distinctive" in the stimuli-lists and because they violate expectations for the schemas activated (Hirshman, 1988; Schmidt, 1991; Hunt & Smith, 1996; Waddill & McDaniel, 1998). This might predict that increases in "bizarreness" would lead to better recall, which is not the case. Items that activate a particular domain-concept but block standard inferences from that concept are not recalled as well as those that maintain such default inferences.

The result suggests that representation of bizarre items does not necessarily end with a check-list in terms of compatibility with activated schemata, but also triggers inferences about possible scenarios that include the object described. The ease with which such inferences are generated may be an important factor in boosting retrieval.[8] This interpretation in terms of inferential potential may explain why most culturally successful religious concepts only use *limited* violations of domain-level information, as illustrated not just in the domain of religious concepts, but also in mythologies, where metamorphoses for instance typically occur between taxonomically close ontological categories (e.g., from person to animal) more often than between distant ones (e.g., from person to artifact) (Kelly & Keil, 1985).

Our experiments focused on the role of recall and inference in cultural transmission. By contrast, cognitive anthropologists interested in religion have often stressed the fact that supernatural agency is "useful" in producing causal explanations of various unexplained events (Whitehouse, 1992; Strauss & Quinn, 1997; Bloch, 1998). In the present view, these two interpretations are congruent. That cultural concepts can be used in causal explanations (and be selected by virtue of their potential for such use) requires that they trigger sufficiently rich inferences. Differential recall and explanatory "usefulness" might be two consequences of the connections established between particular cultural input and universal domain-level expectations.

Notes

1. It is still debated whether the development and adult representations of domain concepts requires prior principles or only abstraction from knowledge of kind-concepts (see Quinn & Johnson, 1997 for an alternative to principle-based scenarios). However, the end-point of the developmental path is clear: Specific information is stored in domain-concepts as opposed to kind or "entry-level" concepts (Jolicoeur et al., 1984; Chumbley, 1986; Kosslyn et al., 1995). Also, some cognitive pathologies display a selective loss of information that affects a whole domain (e.,g. living things) but only impairs kind-level in that domain (e.g., names for living things), see for example, Warrington & Shallice, 1984; Shallice, 1987; Sartori, Coltheart, Miozzo, & Job, 1994; Moss, Tyler, & Jennings, 1997; Kurbat & Farah, 1998.
2. This kind of narrative format is found the world over, for example, in shamans' narratives or pilgrim's accounts, where narrative structure reduces to a list of distinctive places or situations with no causal connections between the episodes.
3. We used the term "pieces of furniture" rather than "objects," "machines" or "artifacts" because the former corresponds to a single, basic-level and frequent lexeme in French ("meuble"). This also blocked possible associations with intelligent machines, supercomputers, etc.
4. For instance, given the stimulus "there were pieces of furniture that could float in the air," neither "there were strange pieces of furniture" nor "there were people that flew around in a room" counted as recall. On the other hand, "there were funny pieces of furniture; they go up rather than down if you let go of them, and people are surprised" was scored as positive.
5. Formal questionnaires are awkward with nonschooled participants in the first place. Also, claiming in this cultural setting that one "knows about" a topic such as witchcraft implies that one is a practitioner.
6. Identifying (linguistically relevant) first language is not always easy in an urban African setting. About half of the population of Libreville have French as their first language, due to frequent marriages between members of different linguistic groups.
7. It is largely debated in Tibet and Nepal whether Bon is a variant of Buddhism or an altogether different doctrine. However, these debates (however politically charged) are academic in the present context. Like Buddhist traditions, Bon teaches that we are

trapped like all other creatures, including gods, in a never-ending cycle of reincarnations as beings who desire and therefore suffer. It is possible to escape from this cycle by realizing how false or illusory "reality" is. On Bon-Po and Buddhist doctrines, the best source is Snellgrove (1959, especially pp. 19–33; see also Ramble 1984).

8. However, it must be noted that spontaneous inferences in sentence-understanding and differences in richness between inferences are not necessarily mysterious phenomena. They are precisely described in pragmatic theories of utterance comprehension (Sperber & Wilson, 1995) and some aspects of this "relevance" approach are experimentally testable (Jorgensen, Miller & Sperber, 1984; Sperber, Cara & Girotto, 1995).

Acknowledgments

Research for this paper was funded by a special grant (#AS41) from the "Cognisciences" Program of the Ministry for Scientific Research, France. The Lyons experiments were conducted at the Laboratoire d'Etude des Mécanismes Cognitifs, Université Lumière, Lyon, with assistance from Professor Nathalie Bedoin and Dr Hervé Bruni. The Gabon studies were conducted at the Ecole Normale Superieure de Libreville with generous support from its Director, Professor Auguste Moussirou-Mouyama as well as from Professor Patrick Mouguiama-Daouda, Université de Libreville. For help with the Katmandu experiments we are grateful to the Abbot of Triten Norbutse monastery, Katmandu for allowing us to run this study, to Vice-Abbot Tenpa Yungdrung for his generous help, and to the monks who participated. Thanks to Harvey Whitehouse, Sheila Walker, Carlo Severi, Michael Houseman, E. Thomas Lawson, Robert McCauley and especially Justin Barrett and Sandra Waxman for comments on a draft version of this paper.

Appendix 1. Literal translation of sample story from Experiment 2

Mr. Wurg was about to be sent as an ambassador to the Zenon 3 galaxy. He wanted to know what things are like over there before leaving. So he went to the Arts and Sciences Museum, where two halls contain exhibits about Zenon 3.

In the first hall, there were exhibits about the various kinds of furniture you can find in houses in the Zenon 3 galaxy. Some of these objects are like what you find here on Earth, and others are really different. There are objects that can be taken apart to be fixed. You can take a part out to change it. There are objects that are aware of what's around them. They know what's going on. There are objects that see what's in front of them. They can perceive what is opposite. There are objects made with parts from other objects. Someone took bits of other objects to make them. There are objects that hide away when they're scared. If something frightens them they run for cover. There are objects designed by engineers. These people made a blueprint of the objects before making them. There are objects that hear sounds around them. They can perceive sounds not too far from them. There are objects made by people as a hobby. They like spending a Sunday making them. There are objects that go where they want to go. If they plan to get somewhere they just go there. There are objects made of a special metal that's hard to melt. It is difficult to make

objects in that metal. There are objects manufactured in small workshops. There are small places where people make them. There are objects that can notice people are staring at them. If someone looks at them they notice it.

In the second hall, there were exhibits about the various kinds of people who live in Zenon 3. Some of these people are very much like us and others are very different. There are people who remember the past. They can recall what happened to them. There are people who are sad when they are alone. They don't like being on their own. There are people you put together with a screwdriver. You screw the parts together to make them. There are people who try to do what they want. If they want to do something they try it. There are people who are made by machines. Special machines turn them out. There are people who can read books. If they open a book they can read it. There are people made of plaster. One uses plaster to make them. There are people who are manufactured when necessary. When you need more you make some. There are people who don't like being bossed around. If you bully them they don't like it. There are people who must be fine-tuned after installation. They must be adjusted after they are put into use. There are people you can fix yourself. If there is a breakdown you can fix them. There are people who understand jokes. They can get what is funny in a joke.

After he had seen all these exhibits, Mr. Wurg went back to the main lobby of the Museum of Arts and Sciences. He had a coffee at the museum cafeteria. He then went back home and cooked his dinner.

Appendix 2. Literal translation of sample story from Experiment 5, person condition

Nyima is a trader who has been to many places. One day, after he had come back from a journey to distant places, he told his friends about the people he had seen in those far-away countries. Some of those people are like you and me and others seem very different.

There was a person who was at two places at same time. He was at one place and at another place at the same moment. There was another person who was visible if you put on the light. You could see him in daylight too. There was another person who could hear past conversations. He could hear the sounds of what people said in the past. There was another person who weighed more than an ox. It was terribly difficult to lift him. There was another person who could guess future events. He knew exactly what was going to happen. There was another person who was taller than a house. He was very tall and impressive. There was another person who could hear what people said. If people talked next to him he heard what they said. There was another person who ran faster than a horse. He went so fast that the horse could not catch up with him. There was another person who went through walls. If he wanted he could walk right through a wall. There was one person who had no shadow. Even in broad daylight, he did not have a shadow. There was another person who could understand jokes. If you told him a joke he laughed. There was another person who could see through a mountain. He could see what was on the other side of the mountain. There was another person who could remember thousands of different names. He recalled the names of thousands of people. There was another person who could see what was in front of him. If you placed an object next to him he saw it. There was another person who could

see villages very far away. He could even see villages many miles away. There was another person who was five feet tall. His height was a little above that of a woman. There was another person who was at one place at a time. If he was somewhere you could see him there. There was another person who has read thousands of books. He has read huge collections of books.

Nyima had finished telling people what he had seen in these far-away places. He told them that after he had traveled to those places he had felt homesick, and that was why he had come back to be with his family.

References

Ackerman, B. P., Spiker, K., & Glickman, I. (1990). Children's sensitivity to topical discontinuity in judging story adequacy. *Developmental Psychology, 26,* 837–844.

Ainsworth-Darnell, K., Shulman, H. G., & Boland, J. E. (1998). Dissociating brain responses to syntactic and samantic anomalies: evidence from Event-Related Potential. *Journal of Memory and Language, 38,* 112–130.

Barrett, J. L. (1996). *Anthropomorphism, Intentional Agents, and Conceptualizing God.* Unpublished PhD dissertation, Cornell University.

Barrett, J. L. (1998). Cognitive constraints on Hindu concepts of the divine. *Journal for the Scientific Study of Religion, 37,* 608–619.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Science, 4* (1), 29–34.

Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a non-natural entity: anthropomorphism in God concepts. *Cognitive Psychology, 31,* 219–247.

Bartlett, F. C. (1932). *Remembering. A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Bergman, E. T., & Roediger, H. L. I. (1999). Can Bartlett's repeated reproduction experiments be replicated? *Memory and Cognition, 27,* 937–947.

Bloch, M. (1992). *Prey into hunter. The politics of religious experience.* Cambridge: Cambridge University Press.

Bloch, M. (1998). *How we think they think. Anthropological approaches to cognition, memory and literacy.* Boulder, CO: Westview Press.

Bloom, P. (1998). Theories of artifact categorization. *Cognition, 66,* 87–93.

Boyer, P. (1990). *Tradition as truth and communication: a cognitive description of traditional discourse.* Cambridge: Cambridge Universiy Press.

Boyer, P. (1992). Explaining religious ideas: outline of a cognitive approach. *Numen, 39,* 27–57.

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

Boyer, P. (1996a). Cognitive limits to conceptual relativity: the limiting-case of religious categories. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity.* Cambridge: Cambridge University Press.

Boyer, P. (1996b). What makes anthropomorphism natural: intuitive ontology and cultural representations. *Journal of the Royal Anthropological Institute (n.s.), 2,* 1–15.

Boyer, P., & Walker, S. J. (1999). Intuitive ontology and cultural input in the acquisition of religious concepts. In K. Rosengren, C. Johnson, & P. Harris (Eds.), *Imagining the impossible: magical, scientific and religious thinking in children.* Cambridge: Cambridge University Press.

Brewer, W. F. (1985). The story schema: universal and culture-specific properties. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language and learning.*

Brewer, W. F., & Lichtenstein, E. H. (1981). Event schemas, story schemas and story grammars. In J. Long & A. Baddeley (Eds.), *Attention and performance 9.* Hillsdale, NJ: Erlbaum.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: The M.I.T. Press.

Child, A. B., & Child, I. L. (1993). *Religion and magic in the life of traditional peoples.* Englewood Cliffs, NJ: Prentice Hall.

Dulaney, S., & Fiske, A. P. (1994). Cultural rituals and obsessive compulsive disorder: is there a common psychological mechanism? *Ethos, 22,* 243–283.

Gerard, A. B., & Mandler, J. M. (1983). Ontological knowledge and sentence anomaly. *Journal of Verbal Learning and Verbal Behavior, 22,* 105–120.

Goldman, R. (1964). *Religious thinking from childhood to adolescence.* London: Routledge & Kegan Paul.

Goody, J. R. (1977). *The domestication of the savage mind.* Cambridge: Cambridge University Press.

Goody, J. R. (1986). *The logic of writing and the organization of society.* Cambridge: Cambridge University Press.

Gopnik, A. (1993). How we know our minds: the illusion of first-person knowledge of intentionality. *Brain and Behavioral Sciences, 16,* 1–14.

Graesser, A. C., Woll, S. B., Kowalski, D. J., & Smith, D. A. (1980). Memory for typical and atypical actions in scripted activities. *Journal of Experimental Psychology (Human Learning and Memory), 6,* 503–515.

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

Hirshman, E. (1988). The expectation-violation effect: paradoxical effects of semantic relatedness. *Journal of Memory and Language, 27,* 40–58.

Houseman, M., & Severi, C. (1998). *Naven and ritual.* Leiden: Brill.

Hunt, R. R. (1995). The subtlety of distinctiveness: what von Restorff really did. *Psychonomic Bulletin and Review, 2,* 105–112.

Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general: the power of distinctiveness in the context of organization. *Memory and Cognition, 24,* 217–225.

Jorgensen, J., Miller, G. A., & Sperber, D. S. (1984). Test of the mention theory of irony. *Journal of Experimental Psychology: General, 113,* 112–120.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: the origins of concepts of living things. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: domain-specificity in cognition and culture.* New York: Cambridge University Press.

Kelly, M., & Keil, F. C. (1985). The more things change. . . metamorphoses and conceptual structure. *Cognitive Science, 9,* 403–406.

Lawson, E. T., & McCauley, R. (1990). *Rethinking religion: connecting culture and cognition.* Cambridge: Cambridge University Press.

Mandler, J., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-Cultural Invariance in Story Recall. *Child Development, 51,* 19–26.

McCabe, A., & Peterson, C. (1990). What makes a narrative memorable? *Applied Psycholinguistics, 11,* 73–82.

Perner, J. (1991). *Understanding the Representational Mind.* Cambridge, MA: The M.I.T. Press.

Polich, J. (1985). Semantic categorization and event-related potentials. *Brain and Language, 26,* 304–321.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory and Cognition, 19,* 523–542.

Spelke, E. S. (1988). The origins of physical knowledge. In L. Weizkrantz (Ed.), *Thought Without Language.* Oxford: Oxford University Press.

Sperber, D. (1985). Anthropology and Psychology. Towards an epidemiology of representations. *Man, 20,* 73–89.

Sperber, D. (1991). The epidemiology of beliefs. In C. Fraser (Ed.), *Psychological Studies of Widespread Beliefs.* Oxford: Oxford University Press.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach.* Oxford: Blackwell.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57,* 31–95.

Sperber, D., & Wilson, D. (1995). *Relevance. Communication and Cognition [2nd edition].* Oxford: Blackwell.

Spiro, M., & D'Andrade, R. G. (1958). A Cross-cultural Study of some Supernatural Beliefs. *American Anthropologist, 60,* 456–466.

Stangor, C., & McMillan, D. (1992). Memory for Expectancy-Congruent and Expectancy-Incongruent Information: A Review of the Social and Social-Developmental Literatures. *Psychological Bulletin, 111,* 42–61.

Strauss, C., & Quinn, N. (1997). *A cognitive theory of cultural meaning.* New York: Cambridge University Press.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall: Differential processing or privileged retrieval? *Memory and Cognition, 26,* 108–120.

Walker, S. J. (1992). Supernatural beliefs, natural kinds, and conceptual structure. *Memory and Cognition, 20,* 655–662.

Watts, F., & Williams, M. (1988). *The Psychology of Religious Knowing.* Cambridge: Cambridge University Press.

Wellmann, H. (1990). *The child's theory of mind.* Cambridge, MA: The M.I.T. Press.

Whitehouse, H. (1992). Memorable religions: Transmission, codification and change in divergent Melanesian contexts. *Man, 27,* 777–797.

Worthen, J. B. (1997). Resiliency of bizarreness effects under varying conditions of verbal and imaginal elaboration and list composition. *Journal of Mental Imagery, 21,* 167–194.

Zelazo, P. D., & Shultz, T. R. (1989). Concepts of Potency and Resistance in Causal Prediction. *Child Development, 60,* 1307–1315. Zelazo, P. D., & Shultz, T. R. (1989). Concepts of Potency and Resistance in Causal Prediction. *Child Development, 60,* 1307–1315.