Contextualizing Counterintuitiveness: How Context Affects Comprehension and Memorability of Counterintuitive Concepts

M. Afzal Upal, Lauren O. Gonce, Ryan D. Tweney, D. Jason Slone

Abstract

A number of anthropologists have argued that religious concepts are minimally counterintuitive and that this gives them mnemonic advantages. This paper addresses the question of why people have the memory architecture that results in such concepts being more memorable than other types of concepts by pointing out the benefits of a memory structure that leads to better recall for minimally counterintuitive concepts and by showing how such benefits emerge in the real-time processing of comprehending narratives such as folk tales. This model suggests that memorability is not an inherent property of a concept; rather it is a property of the concept, the context in which the concept is presented, and the background knowledge that the comprehendor possesses about the concept. The model predicts how memorability of a concept should change if the context containing the concept were changed. The paper also presents the results of experiments carried out to test these predictions.

Keywords: Culture; Memory; Concepts; Language understanding; Religion

1. Introduction

Several of the factors influencing the individual observer are social in origin and character. In the actual remembering of daily life, the importance of these social factors is greatly intensified. The form, which a rumor, or a story, or a decorative design, finally assumes within a given social group is the result of many different successive social reactions. Elements of culture, or cultural complexes, pass from person to person within a group, or from group to group, eventually reaching a thoroughly conventionalized form, and may take an established place in the general mass of culture possessed by a specific group. In this way, cultural characters, which have a common origin may come to have apparently the most diverse forms. (Page 118, Bartlett 1932)
Sir Frederic Bartlett (1886–1969) was one of the first to systematically study how concepts embedded in stories are transformed as they pass from person to person. In a seminal series of studies using primarily the “Method of Repeated Reproduction” and the “Method of Serial Reproduction,” he compared the distortions produced in memory and in the retelling of narrative stories, examining the processes of omission, distortion, and transformation that occurred (Bartlett, 1932). For example, he asked British university students to read passages from various folk tales, including the Native North American folk tale “The War of the Ghosts” (Erdoes & Ortiz, 1984), and then to retell it to others in writing who then retold it to others. Over successive retellings, Bartlett noted that culturally unfamiliar concepts were often replaced by more familiar concepts; for instance, a canoe was replaced by a rowboat. In none of the series of ten reproductions of “The War of Ghosts,” did a mention of ghosts remain, even though the story’s title mentions ghosts. Bartlett argued that a process of “conventionalization”—a process by which unfamiliar material becomes assimilated to a more familiar schema—was responsible. In effect, he emphasized the powerful role of context on the constructive and reconstructive processes of both memory and cultural transmission. He observed similar processes for both culturally familiar and culturally unfamiliar stories (comparing, for example, a story about a cricket game to unfamiliar folk tales) and in several cultural groups (British university students vs. Indian graduate students).

Bartlett’s findings have been important to recent investigators interested in exploring the roots of religious cognition and understanding why some religious beliefs are more widespread than others (Boyer, 1994; Barrett & Nyhoff, 2001; Boyer & Ramble, 2001; Atran, 2002; Atran & Norenzayan, 2004; Upal, 2005; Gonce et al., 2006). Most of this work, however, has ignored the dynamic character of Bartlett’s theory. For him, “memory” was a dynamic process in which comprehension, existing schemas, and current needs and interests interacted to reconstruct remembered events. By contrast, most recent research has been “item-centric,” as if recall of a specific item depended only upon properties of that item. Barrett and Nyhoff (2001), for example, used six Native North American folk tales from Erdoes and Ortiz (1984) in a recall task. The stories of about 500 words or less were chosen so that they contained both intuitive concepts such as river, mountain, and bird, and expectation violating counterintuitive concepts such as a talking bird and a walking stone. Unlike Bartlett, they found that the recall rates for counterintuitive concepts were significantly higher than the recall rates for intuitive concepts. Barrett and Nyhoff (2001) also designed an artificial story to better control for the number of intuitive and counterintuitive concepts, narrative structure, and the amount of repeated exposure to a concept. The futuristic story about a person visiting a museum to see alien beings and artifacts was designed to contain six concepts of each of the following three types:

- intuitive concepts such as a being who is aware of its existence,
- counterintuitive concepts such as a being who never dies, and
- bizarre concepts such as a being who weighs 1000 pounds.

Using Bartlett’s method of serial reproduction, they found that after three retellings, counterintuitive concepts were better recalled than bizarre concepts, which were better recalled than intuitive concepts.¹
Boyer and Ramble (2001) used a variant of Barrett and Nyhof’s (2001) alien museum story but did not use a serial reproduction task. Instead, they had subjects read the story and, following a brief distraction task, answer a question requiring reproduction of the intuitive, counterintuitive, and bizarre items mentioned in the story. Their results supported Barrett and Nyhof’s findings that counterintuitive items are best recalled and that intuitive items are worst recalled. Across six studies, they found similar effects in three different cultural settings, in France, Nepal, and Gabon. Further, they found that the extent to which expectation violations affected recall depended upon whether the violations occurred at the domain level or at the kind level, and whether they involved violations of specific expectations (“breaches”) or transfers of inappropriate properties (“transfers”). They suggested that “distinctiveness effects” (Waddill & McDaniel, 1998) could be responsible for the observed results, in that distinctiveness could represent an adaptive “cognitive optimality” for recall.

Atran (2002) and Atran and Norenzayan (2004) argued that basic concepts in folk psychology, folk biology, and folk physics include expectations about a limited number of ontological categories. Supernatural concepts typically involve violation of only one of these basic ontological categories. Atran and Norenzayan constructed three lists of intuitive concepts, minimally counterintuitive concepts (concepts that violate only one ontological expectation such as a sobbing oak), and maximally counterintuitive concepts (concepts that violate two ontological expectations such as a squinting wilting brick). Subjects were presented lists of concepts without the narrative structure used by previous researchers. Each subject saw a list containing an equal number of all three types of concepts and was asked to recall as many items as he/she could after a brief distraction task. Like Bartlett, they found that subjects recalled intuitive concepts better than minimally counterintuitive concepts which were better recalled than maximally counterintuitive concepts. Removing the narrative structure used by Barrett and Nyhoff (2001) and Boyer and Ramble (2001) thus appeared to lower the recall rates for counterintuitive concepts. Atran and Norenzayan (2003) suggested that recall for counterintuitive concepts increases when they are ‘mixed in’ with a larger number of intuitive concepts. This is in accordance with the results from a number of studies of visual stimuli that report an enhanced recall for distinct and bizarre stimuli when surrounded by mundane stimuli (Waddill & McDaniel, 1998). While the mnemonic advantage of bizarre items has long been acknowledged (and plays an important role in memory enhancement techniques, such as the “Method of Loci”), much recent research suggests that the effects of bizarreness on memory are not simple (McDaniel, Einstein, Delosh, & May, 1995). Thus, Worthen and Roark (2002) showed that bizarre material was better recalled than common material, but that it was also more subject to distortion and merging with other memory fragments. McDaniel, Dornburg, and Guynn (2005), using mixed and unmixed lists of bizarre and common sentences, showed that bizarreness effects are mediated by retrieval dynamics, not by encoding processes, suggesting that such effects are one aspect of a more general distinctiveness effect.

Gonce, Upal, Slone, and Tweney (2006) examined recall for intuitive (INT), counterintuitive (MCI), and maximally counterintuitive (MXCI) items in both lists and in story-like contexts. In both conditions, INT and MCI items were recalled better than MXCI items. The differences between INT and MCI recall were slight, however, and, by manipulating the immediate
context, could be reversed in either direction. This result strongly suggests that an “item-centric” explanation of recall effects in such situations is not possible. Instead, the interaction between items and context must be taken into account. Of particular importance for the present study, note that results of Gonce et al., while mirroring the phenomena usually discussed under the heading of bizarreness and distinctiveness, were obtained in prose-like contexts of connected narrative. That is, rather than an effect limited to items in lists, the focus was upon the effect of context on items within a holistic text.

In the present paper, we outline a theory that explains why minimally counterintuitive concepts are better remembered and recalled compared to both intuitive and maximally counterintuitive concepts. Like Bartlett’s, our account emphasizes the dynamic processing of material to be remembered, although our theory is framed in the context of recent work in AI and in cognitive science.

We start with the analysis of the memory architecture of an abstract problem solving agent2. We believe that this type of analysis is useful as it points out the evolutionary advantages conferred to agents with a memory architecture that exhibits specific characteristics. This approach permits us to derive predictions about relative recall rates and also provides some insight into the kind of concepts that are likely to be remembered and recalled by people. Note that such an analysis presumes that an adapted architecture may possess an advantage in an environment in which natural selection is operating. However, the analysis given here presumes only that the adaptive advantage is relatively short-term and can be framed in, for example, cost-benefit terms.

2. Memory for intuitive and counterintuitive concepts

Memory is critical to the survival of problem solving agents that are frequently called on to make accurate predictions about their environment in order to secure resources, to safely negotiate their way, and to avoid falling prey to various predators. While some have advanced alternative architectures (Brooks, 1991), most cognitive scientists believe that this requires such agents to build internal representations of their environment (the so called world model). When confronted with a novel object/situation, an agent with limited cognitive resources has to weigh the potential savings that may be obtained by remembering the new situation/object and recalling it when needed against the costs of having to generate that information using its existing world model. If the cost of producing a piece of information using an agent’s existing knowledge about the world is less than the cost of remembering and recalling it, then that piece of information should get lower priority for memorizing3. On the other hand, if it is computationally harder to generate the information using one’s existing knowledge than to remember and recall it, then that piece of information should get preferential access to memory resources. In order to make this tradeoff an agent needs to know three things:

1. The likelihood that the given piece of information will be needed in the limited time horizon in the future (i.e., within the time horizon by which it is not likely to have been deleted or become inaccessible from memory), $L$. 

2. The cost of producing the information using agent’s current knowledge base, \( \text{Cost}_{\text{Production}} \), and
3. The cost of storing information and retrieving it from memory, \( \text{Cost}_{\text{Storage + Retrieval}} \).

Evolution should bias the memory of an intelligent agent, whose survival depends on being able to accurately predict information about its environment, to preferentially remember and recall those pieces of information that maximize \( L \times (\text{Cost}_{\text{Production}} - \text{Cost}_{\text{Storage + Retrieval}}) \). However, computing these quantities directly is computationally burdensome. For instance, one way to compute \( \text{Cost}_{\text{Production}} \) is to use the current knowledge base to actively generate predictions about the future. Indeed Schank (1979) and others (e.g., Kintsch, 1980) have argued that people actively generate such expectations to test the effectiveness of their world model at predicting real world happenings although there must obviously be limits to the amount of such testing that can be carried out. If these expectations are confirmed by real world events then belief in the validity of the world model is strengthened. If, on the other hand, the expectations are violated, then people realize that their current world model is faulty and attempt to revise it, for instance, by remembering the new episode as a distinct new case unrelated to previous cases (Schank & Abelson, 1977). Schank and others have argued that people engage in this expectation-based reasoning strategy when reading to comprehend stories.

Over the last few decades, however, discourse analysis studies have accumulated overwhelming evidence that indicates that when reading to comprehend stories people primarily attempt to answer the ‘why’ questions as opposed to answering the ‘what happens next,’ or ‘how,’ and ‘where’ questions (Graesser et al., 1997). That is, people attempt to find justifications to explain why an author mentions a particular piece of information in the text. These explanations help them to integrate disparate textual units into a coherent higher level representation. The highest level representation is an overall coherent theme for the story. The more disparate a textual unit, the more cognitive effort is required to process it. Kintsch (1980) defines postdictability as the ability of a text to “hang together and make sense to the user, so that he is able to construct a coherent macrostructure in which each text unit has its place and is meaningfully related to other sections of the text.” (p. 89). In this paper, we use the term postdictability to refer to the ease with which a concept’s inclusion in the text can be justified after the textual unit containing that concept has been read. This can be contrasted with the predictability of a textual unit as the ease with which the occurrence of the concept can be predicted prior to the concept having been read. Ease of constructing a justification or prediction is inversely related to the computational effort required to construct a justification or prediction.

Clearly, predictability and postdictability are not independent of each other. In particular, as a text is read, what was prediction at one moment will soon be part of the basis of a postdiction and a further prediction. While the two inferential processes are thus dynamically changing, it is possible to define them separately, especially for relatively short texts. We define the prior context of a target concept as the text that precedes the occurrence of the concept and the posterior context of a target concept as the text that immediately follows that concept. For instance, consider the short narrative shown in Fig. 1. The title, place, date and the first sentence form the prior context in which the concept, “the cow flying above the trees” occurs.
Prediction and postdiction seem to require two different types of reasoning processes. The ability to predict what comes next seems to require the ability to generate expectations about the future using a generative process similar to problem solving and planning (Newell & Simon, 1972) while the ability to find justifications seems to imply an abductive process (Ng & Mooney, 1990). In most situations, predicting an outcome before its occurrence is more computationally challenging than justifying an outcome after it has happened. This is because in most common sense reasoning situations, the space of possible worlds one has to search starting from a given state of the world to find the state most likely to occur next is larger than the space of explanations one has to search through to find a justification for an outcome that is already known to have occurred.6

Prediction-based theories have been especially popular in artificial intelligence and machine learning (Schank, 1979; Ram, 1987; Redmond, 1992) because they allow an agent to recognize a learning opportunity. When an expectation fails, an agent can recognize the inability of its current world model and/or inference processes to make correct predictions. It can use this opportunity to revise its knowledge and/or change its inference processes. However, justifications constructed through postdiction can also be used to make predictions about the future. Indeed, there is some evidence to indicate that that may be how people make predictions as they read text (Smith & Hancox, 2001). Most models of comprehension (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992) concede that a reader may have to abandon a justification as more and more evidence accumulates against it and thus search for an alternative justification. The decision to abandon a justification can also indicate a shortcoming in the reader’s knowledge or inference process (similar to a prediction-failure). This may suggest that the reader is dealing with an unusual or novel entity, or event.

While all language comprehension models assume that readers attempt to derive some form of coherence among various components of the text they read,7 they assume that meanings of various concepts constituting a sentence or a paragraph are known. However, as Franks (2003)8 points out many religious concepts are in fact conceptual combinations and in order to understand how texts containing them are understood, we cannot ignore how meanings of such conceptual combinations are derived. Although there is little agreement as to the
precise form in which concepts should be represented, most researchers agree that concepts are connected to people’s commonsense notions about that domain. Combined concepts are formed by either transferring properties from one concept to the other or by relating the two concepts to each other through a relationship that is not a part of either of the concepts to be combined. It is now agreed by most that conceptual combination is a knowledge-driven process in which a subject’s background knowledge of the properties of the constituent concepts plays a critical role. However, it is not well understood as to precisely what that role is, how dynamic changes in a subject’s knowledge about a concept brought about during the reading of a text impact the concept and how those changes in turn impact the computation of local and global coherence.

Our account of the reciprocal roles of predictability and postdictability resembles in some respects the pragmatic “Relevance Theory” developed by Sperber and Wilson (1995). As in our account, Sperber and Wilson argued that inputs to the comprehension system vary in relevance to a listener, and that degree of relevance determines what meaning the listener will derive from an utterance. The most relevant interpretation of an utterance is the interpretation that activates the largest amount of listener’s prior knowledge while requiring the least amount of cognitive effort. Van der Henst and Sperber (2004) tested predictions drawn from this theory using relational reasoning tasks, Wason four card selection tasks, and a field study in which people were asked for the time and the extent of numerical rounding was measured. In all three situations, relevance principles predicted the kinds of responses observed. Relevance Theory thus appears applicable to the understanding of discourse processes, and draws upon optimality principles that resemble those in our account. In this sense, our account can be seen as an extension of Relevance Theory to the domain of memory for contextually embedded items. Given an utterance, while Relevance Theory explains why listeners derive the meanings that they do (because they are the most relevant to them), our account explains why listeners remember the concepts that they do (because they provide the most information gain).

Note also that our account is similar to the *adaptive rationality* approach of Anderson (1990; see also the papers in Oaksford & Chater, 1998). Like Anderson, we model a comprehendor based upon presumed calculations of likelihood, weighted by cost considerations. Similarly, Gigerenzer and his colleagues (e.g., Gigerenzer, Todd, & The ABC Group, 1999) have argued that simplifying heuristics that reduce cognitive computational load are adaptive features of much cognition. For example, in one study stock market picks that were made using only one feature, the “recognizability” of the stock name, outperformed a group of stock market analysts in picking winners. Gigerenzer et al., argued that such “Fast and Frugal Heuristics” are especially adaptive in probabilistic environments characterized by much uncertainty. For the specific case, picking “Coca Cola” is no sure way to market success, but the recognizability of the name is likely to correlate with successful performance, which is why it is recognizable in the first place. In this sense, the “Minimal Counterintuitiveness Hypothesis” may represent a similar adaptive heuristic.

2.1. The memorability hypothesis

We use the notions of predictability and postdictability as defined above to restate our intuitions about the types of narrative concepts that an adaptive agent should preferentially
Fig. 2. The variation in the space of concepts along the dimensions of predictability, postdictability and memorability. Intuitive (INT) concepts have high predictability and postdictability, maximally counterintuitive (MXCI) concepts have low predictability and postdictability while minimally counterintuitive concepts (MCI) have low predictability but high postdictability.

remember. Predictability of a concept is inversely related to the cost of having to create a textual unit: a high predictability concept has low cost of creation whereas a low predictability concept is harder to create. Since the explanation that postdiction produces can be used by an agent to diagnose problems with its existing knowledge and modify it, postdictability of a concept indirectly provides an estimate of the ease with which a novel observation can be integrated into an agent's existing world knowledge. Thus a rational agent should preferentially remember those concepts that have a high postdictability and a low predictability. That is, memorability of a concept $C$ is:

$$\text{Memorability}(C) = K \times L(C) \times (\text{postdictability} - \text{predictability}).$$

Here, $K$ is a constant and $L(C)$ is the likelihood that $C$ would be encountered in the near future. This hypothesis explains why intuitive concepts may have low memorability—because they have high predictability and postdictability—while maximally counterintuitive concepts are less memorable because they have low predictability and postdictability. As shown in Fig. 2, minimally counterintuitive concepts achieve a cognitive optimum because they lie in the region of high postdictability and low predictability values.

An important contribution of the memorability hypothesis is that it allows us to understand that there is nothing inherent or magical about minimally counterintuitive concepts. Rather than a property of the item itself, as most previous researchers have assumed, it is the properties of the item in a particular context that matter. Here, counterintuitiveness is regarded as a relational property of the item-in-a-context, and, within our model, a computationally-derived parameter. Minimally counterintuitive concepts have a higher $(\text{postdictability} - \text{predictability})$ value and higher memorability than intuitive and maximally counterintuitive concepts in certain contexts and lower $(\text{postdictability} - \text{predictability})$ value and lower memorability in other contexts. Indeed, our hypothesis predicts that memorability for all types of concepts can be increased or decreased by varying their postdictability and predictability values. This can be done by:

- changing the verbal context in which a concept appears (Gonce et al., 2006, in press). This is the option that we pursue further in this paper;
by not giving subjects enough time to generate elaborations required to justify the occurrence of a concept in the given context (Heit, 1998) e.g., by presenting the material for short period of time; or
by varying the instructions given to subjects to process the concept and its context (e.g., “explain what this means” versus “why does it make sense” versus “read aloud once”) (O’Reilly et al., 1998).

Our hypothesis predicts that if all else remains the same and

1. if a concept is made more predictable (e.g., by changing the prior context) then its memorability should decrease,
2. if a concept is made less predictable (e.g., by changing the prior context) then its memorability should increase,
3. if a concept is made more postdictable (e.g., by changing the posterior and the prior context but without affecting its predictability) then its memorability should increase, or
4. if a concept is made less postdictable (e.g., by changing the posterior and the prior context but without affecting its predictability) then its memorability should decrease.

3. Evaluating the memorability hypothesis

The main difficulty in evaluating the memorability hypothesis is that of measuring changes in postdictability and predictability on the same scale so that they can be compared. Since both are measures of computational effort, such a measurement is possible in principle but it requires a detailed computational model of the cognitive processes underlying comprehension of narrative texts, something that we currently do not have. However, an alternate approach is to attempt to vary only one of these quantities (i.e., postdictability and predictability) while keeping the other one constant and measuring the changes in memorability to see whether those changes happen as predicted by our hypothesis. This is the approach that we adopt here.

We designed two studies to test two hypotheses:

(A) postdictability values of intuitive (INT), minimally counterintuitive (MCI), and maximally counterintuitive (MXCI) concepts should differ significantly from each other, in the order postdictability (MXCI) < postdictability (MCI) < postdictability (INT);
(B) changes in context should impact concept memorability as predicted by the memorability hypothesis.

Note that due to the inability to compare changes in predictability with changes in postdictability, we believed that nothing could be gained by using the same stimuli for both experiments. For our first study, designed to test hypothesis A, we used lists of items similar to those used by Atran and Norenzayan (2005) to see if differences in postdictability values of these concepts can explain the observed differences in recall. The second experiment, designed to test hypothesis B, modified the stories used by Barrett and Nyhoff (2001) to vary predictability of counterintuitive concepts by changing the prior context of these concepts.
Using the materials used by previous researchers allowed us to confirm aspects of the findings of these researchers and to use their results as a baseline for our experiments.

3.1. Experiment 1

The first experiment used three types of items used by Norenzayan and Atran (2005): Maximally counterintuitive (MXCI), minimally counterintuitive (MCI), and intuitive (INT). We corrected a potential confound problem in Norenzayan and Atran’s study by adding two words to each intuitive concept and one word to each minimally counterintuitive concept to make all concept names three words long. This should eliminate potential differences in recall due to item length. Two types of posterior context were added to obtain items with context: supportive context attempted to make concepts more believable (by increasing their postdictability), and contradictory context attempted to make concepts less believable (by decreasing their postdictability). Two versions of MCI concepts with context were designed: S-MCI are MCI concepts with supportive context added to them and C-MCI are MCI concepts with contradictory context added to them. For instance, Fig. 3 shows the MCI concept, ‘flying healthy cow,’ with contradictory and supportive context added. Only supportive context was added to the MXCI concepts to increase their postdictability. Two versions of INT concepts were designed: C-INT items were intuitive concepts with contradictory context added to make them less believable, and S-INT items were intuitive concepts with supportive posterior context.

To assess postdictability, we used a rating procedure which asked subjects to assess each item’s meaningfulness, by asking, for example, “What could a ‘flying healthy cow’ mean?” They were then asked to assess the difficulty of answering this question. Note that this rating does not directly measure the processes by which justification occurs, nor does it tap the nature of the conceptual representation achieved by the subject. Instead, it related to the subject’s own assessment of the cognitive load that constructing justification might require. Since assessment of cognitive load is part of what our model requires of subjects, these ratings should be correlated with the overall ease or difficulty of justification. In effect, we are asking subjects to judge the relevance of representations, rather than the representational process itself. That is, our subjects must assess the likely relevance of a representation to determine its

<table>
<thead>
<tr>
<th>Contradictory Context</th>
<th>Supportive Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flying healthy cow. Flying cow is an example of the empty set, said Professor Pythagoras. Cows cannot fly or even jump very high in the air. Cows are very heavy animals and they do not have strong leg muscles required to jump high like Michael Jordan.</td>
<td>Flying healthy cow. The old Lapp goddess Mittshwafen is also known as the ‘flying cow’ because of her miraculous ability to fly in the air bestowed by the Lapp’s patron Saint, St. Arthur, after the cow’s generosity in offering the poor free milk during the famine of 1429.</td>
</tr>
</tbody>
</table>

Fig. 3. A minimally counterintuitive (MCI) concept with contradictory and supportive posterior context added.
difficulty. Such judgments require processing at the pragmatic level, as in some of the tasks used by Van der Henst and Sperber (2004).

3.1.1. Participants

Participants were 63 University of Findlay undergraduate students. Participation was part of a classroom demonstration and students were given extra course credit for their participation.

3.1.2. Materials and procedure

Subjects were asked to rate the concepts by how hard or easy it was for them to justify each concept. Each packet contained eight items; one item of each one of the eight types of INT, C-INT, S-INT, MCI, C-MCI, S-MCI, MXCI, and S-MXCI. Half of the items were presented without a context while the other half had a posterior context added to vary their postdictability (note that there are no C-MXCI items). Items with and without context were presented in an alternating order: each item without context (except the last item) was followed by a context item and each context item (except the last item) was followed by an item without context. Half of the subjects saw an item without a context as the first item while the other half saw an item with a context as the first item. Ten packages were designed by varying the order in which items appeared in the packet. Appendix A shows part of a packet used in Experiment 1.

3.1.3. Results and discussion

The subject ratings (between 1 and 10) of the difficulty of justifying a concept were subtracted from 10 to compute the postdictability, i.e. the ease of justifying a concept. Figure 4 shows the mean postdictability ratings for the five items with context. The overall effect of item type was significant, $F(4, 270) = 7.152$, $p < 0.001$. All of the means fall in the predicted order of recall, although the key comparisons, S-MCI > C-MCI and S-INT > C-INT, were not significant when Bonferroni-corrected post hoc tests were used, $p = 0.268$ and $p = 0.666$, respectively. The extreme comparisons, C-MCI versus C-INT and S-INT and S-MXCI vs. S-INT were significant, both $p < 0.05$.

The mean postdictability ratings for the three items presented without context is shown in Fig. 5. As predicted by the memorability hypothesis, there was an overall effect of item type $F(2, 274) = 29.086$, $p < 0.001$. Postdictability values of INT concepts were significantly higher than the postdictability values of MCI concepts [$postdictability(MCI) < postdictability(INT)$, $p < 0.001$] which were in turn significantly higher than the postdictability ratings of MXCI concepts [$postdictability(MXCI) < postdictability(MCI)$, $F(1, 163) = 5.419$, $p = 0.032$]. This confirms Hypothesis A.

The differences in postdictability ratings are consistent with findings of the recall experiments originally performed by Atran and Norenzayan (2005) and repeated by Gonce et al., (2006). This is because in these experiments, subjects saw lists containing equal numbers of various types of concepts, without a coherent context, and in no particular order. This means that subjects were not able to form expectations about the kind of concepts (i.e., INT or MCI or MXCI) to follow. In other words, for concepts embedded in lists containing equal numbers of different types of items presented randomly, there are no large differences in predictability values, i.e.,
Predictability(MXCI) \approx \text{predictability(MCI)} \approx \text{predictability(INT)}

Thus memorability of concepts in such lists depends on postdictability values alone. Since subjects assigned highest postdictability values to intuitive concepts, the memorability hypothesis predicts that intuitive concepts should have the best memorability. Maximally counterintuitive concepts should have the lowest memorability because they have the lowest postdictability values. This is exactly what Atran and Norenzayan (2004) and Gonce et al. (2006) observed (as shown in Fig. 6) in their recall experiments with lists containing an equal number of INT, MCI, and MXCI items.

### 3.2. Experiment 2

In this set of experiments, we measured and varied concept predictability to explore its impact on memorability of intuitive and minimally counterintuitive concepts. We used variants of two stories used by Barrett and Nyhoff (2001), namely the alien museum visit story (also used by Boyer & Ramble, 2001) and a story about the journey of a brother and sister from school to home. The alien museum story, “The Adventures of Mr. Wurg” was rewritten to contain six counterintuitive and six intuitive concepts (See Appendix B). “The Journey Home/My Dream” story contained six counterintuitive concepts and twelve intuitive concepts (See Appendix C). Two versions of the stories were created: in the counterintuitive-supportive
version, the prior context of the counterintuitive concepts was modified to make them more predictable while the intuitive-supportive version used the prior context used by Barrett and Nyhoff (2001). The counterintuitive-supportive versions of both stories also modified the opening paragraphs of both stories to prepare the reader to expect some of the counterintuitive concepts to follow (as shown in Fig. 7). For instance, the opening paragraph of the journey home story was modified to add that the narrative to be followed was a boy’s dream and that the boy had a history of dreaming about things divorced from reality. The title of the story was also changed to “My Dream.” The opening paragraph of “Adventures of Mr. Wurg” was changed to state that the alien galaxy possessed technology far advanced from that of earth resulting in incredible inventions such as intelligent robots who cook for their owners.

3.2.1. Experiment 2A

This experiment was designed to test the hypothesis that subjects will be more likely to expect counterintuitive concepts in the counterintuitive-supportive version of the stories than in the intuitive-supportive version of the concept, i.e. subjective predictability ratings for MCI-concepts embedded in counterintuitive-supportive concepts will be significantly higher.
Fig. 6. A summary of mean recall rates observed by Gonce et al. in their experiments with INT, MCI, and MXCI items presented in lists without context. Error bars show ±1 standard error of the mean.

<table>
<thead>
<tr>
<th>Prior context added to “The Adventures of Mr. Wurg”</th>
<th>Prior Context added to “The Journey Home—My Dream”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Razonians love smart object technology. Most Razonian objects have digital sensors and artificial intelligence chips embedded in them to allow them to perceive their environment and act intelligently.</td>
<td>I have always been fascinated by dreams. I have always wondered as to why some of our dreams are so different from our everyday experience; why are the laws of nature violated so often in our dreams even though that never happens when we are awake. I remember that in some of my childhood dreams, our childhood puppy Jack would talk to me. In others, I would fly through the air.</td>
</tr>
</tbody>
</table>

Fig. 7. The prior context added to the stories to make the counterintuitive ideas to follow more predictable and thereby decrease their memorability.
than the predictability ratings for MCI concepts embedded in intuitive-supportive versions of the stories.

Predictability was assessed by asking subjects, prior to reading the story but after reading the preamble, to rate the likelihood of occurrence of the concept in the following story. Note that this rating is a direct reflection of the predictability of the concept, i.e. the ease with which the reader can predict the occurrence of a concept to follow.

3.2.1.1. Participants: Twenty nine undergraduate psychology students from Bowling Green State University participated in the study as a part of their undergraduate course work.

3.2.1.2. Materials and procedure: Two types of packets were created: ‘intuitive concept supportive packets’—packets with the original prior context used by Barrett et al. (2001) and Boyer et al. (2001) and ‘counterintuitive concept supportive packets’—packets with modified prior context intended to increase predictability of counterintuitive concepts by prompting subjects to expect counterintuitive concepts in the story to follow. Each packet contained an intuitive or counterintuitive concept-supporting preamble (including the story title and/or introductory sentences) followed by a list containing a short description of each concept. The subjects were asked to rate the concepts (on a scale from —2 to 2) by their likelihood of occurrence in the given story. Half of the subjects (14) received the original intuitive-supportive context stories while the other half received stories with counterintuitive-supportive context.

3.2.1.3. Results and discussion: The mean predictability ratings for the counterintuitive concepts embedded in both stories is shown in Fig. 8. They show that addition of counterintuitive-supportive prior to the context was successful in increasing the predictability of counterintuitive concepts. Note that there was a strong main effect due to story, F(1,27) = 47.641, p < 0.001, with “Adventures of Mr. Wurg” being less predictable, p < 0.001, as well as a main effect due to the kind of context, F(1,27) = 4.915, p = 0.035 with counterintuitive-supportive context leading to greater predictability, p < 0.001!. The interaction was also significant, F(1,27) = 8.163, p = 0.008.

Having made effective changes in the stories to manipulate predictability we wanted to see if these changes had an impact on the recall of counterintuitive concepts. Experiment 2B tested this question.

3.2.2. Experiment 2B

This experiment was designed to test the hypothesis that higher predictability for MCI concepts in counterintuitive-supportive versions of the stories (as compared to the predictability of counterintuitive-concepts in intuitive-supportive versions) should result in significantly lower recall rates for counterintuitive concepts in counterintuitive-supportive stories than recall rates of counterintuitive concepts in intuitive-supportive stories.

3.2.2.1. Participants: Eighty-four University of Findlay students ranging in age from 18 to 35 years, 40 male, and 44 female with a mean age of 21 years participated in the study as a part of their undergraduate course work.
3.2.2.2. **Materials and procedure:** Packets each containing either counterintuitive-supportive versions of both stories or intuitive-supportive versions of both stories were designed. Four versions of the “Adventures of Mr. Wurg” were designed by changing the order in which Mr. Wurg observes museum objects to avoid order effects. Since changing the order of events in “the journey home/my dream” story would have broken the plot of the story, a single version of that story was used. The order of the two stories was counterbalanced so half the subjects read “Adventures of Mr. Wurg” first while the other half read “the journey home/my dream” story first. Half of the students (42) were randomly selected to receive the counterintuitive-supportive version of both stories while the other half received the intuitive-supportive versions. Subjects were instructed to carefully read each story trying to imagine a situation as they read it. Once all subjects had read a story, they were given a distraction task involving simple arithmetic for the next three minutes. Following the distraction task, they were asked to write down as many of the items from the story as they could recall.

The subjects then repeated the process of reading a story, performing a distraction task for three minutes, and writing down all the items, for the second story. At the end of the session subjects were told about the aims of the study.
Fig. 9. Mean proportion of intuitive (INT) and counterintuitive (CI) concepts recalled by subjects who read the original intuitive-supported version and subjects who read the modified counterintuitive-supportive versions of “Adventures of Mr. Wurg” and “The Journey Home/My Dream.” Error bars show ±1 standard error of the mean.

3.2.2.3. Results and discussion: Three responses were discarded because the subjects were unable to complete the task. The remaining 81 responses were scored by a hypothesis-blind coder. Six counterintuitive and twelve distinct intuitive concepts were identified in the original “journey home/my dream” story. The coder compared the items recalled by the subjects with the items given in the story and counted the number of intuitive and counterintuitive concepts faithfully recalled by the subjects. This number was then divided by the total number of intuitive and counterintuitive concepts present in the original story to measure the proportion of intuitive and counterintuitive concepts recalled by each participant.

As in the studies by Boyer and Ramble (2001) and Barrett and Nyhoff (2001), subjects who read the original intuitive supportive versions of the two stories recalled a significantly larger proportion of counterintuitive concepts than intuitive concepts, F(1,79) = 47.980, p < 0.001. There was a main effect of story, F(1,79) = 12.856, p = 0.001, with higher recall for “Journey/Dream.” There was also a significant effect of context, F(1,79) = 5.494, p = 0.022, with intuitive-supportive versions having higher recall. Subjects who read the counterintuitive supportive versions of the two stories recalled fewer counterintuitive concepts than
subjects who read the intuitive supported versions, though the difference was not statistically significant, $p = 0.098$. Further, recall rates of counterintuitive concepts in counterintuitive-supportive versions of the stories were lower than recall rates for counterintuitive concepts in intuitive-supportive versions. The subjects who read counterintuitive supportive versions of the two stories recalled significantly fewer intuitive concepts ($M = 0.522$) than the subjects who read the intuitive supportive versions of the two stories ($M = 0.580$). The differences between recall rates of intuitive concepts for counterintuitive-supportive and intuitive supportive versions were significant for both stories. This may indicate that the changes in the prior context, though only intended to increase the predictability of the counterintuitive concepts by prompting the user to expect counterintuitive concepts, also affected the intuitive concepts. The changes may have made intuitive concepts harder to justify (e.g., the subjects may have found it more difficult to explain the occurrence of ordinary objects such as, “object made by people as a hobby” in a story about objects belonging to a high tech civilization than they did in the original version of, “the adventures of Mr. Wurg”) thereby decreasing the postdictability of intuitive concepts and lowering memorability for these concepts in the counterintuitive-supportive version. This change in prior context may also have decreased the predictability of intuitive concepts somewhat but the memorability hypothesis predicts that if decrease in postdictability was larger than the decrease in predictability then the memorability for concepts embedded in the counterintuitive-supportive context should be lower. In principle, this can be tested by measuring changes in predictability and postdictability of intuitive concepts that result from varying the prior context. However, as discussed earlier, measuring predictability and postdictability on the same scale to be able to compare them is difficult.

Another result we did not anticipate when designing the experiment is that there was no statistically significant difference in the proportion of intuitive and counterintuitive concepts recalled by the subjects who read the intuitive-supportive version of “adventures of Mr. Wurg.” This may mean that the counterintuitive concepts we embedded in that story were not as “counterintuitive” as those used by Barrett and Nyhoff (2001) and Boyer and Ramble (2001). This is especially noteworthy since the results regarding the impact of context on recall are robust and appeared consistently across all conditions.

The results of the above two experiments strongly suggest that the context in which concepts appear has a significant impact on recall of those concepts. This role is not accounted for by previous theories that attempt to explain better recall for counterintuitive concepts as an intrinsic property of such concepts. Instead, the proposed model explains at least some of the difference in recall as being due to the differences in predictability and postdictability of intuitive and counterintuitive concepts.

4. Conclusion

An analysis of memory in an abstract problem-solving agent seeking to improve mental models of its environment to improve its fitness, allowed us to formulate a computational model of memory. This model specifies that memory of an intelligent agent should evolve to preferentially remember and recall those events/objects that violate the agent’s expectations about the future but can be justified once they have been observed. Minimally counterintuitive
concepts such as religious concepts of ghosts and gods, when they appear in myths and folk tales, meet these requirements. Anthropologists have long argued that this gives such concepts mnemic advantages, thereby allowing them to spread in populations (Boyer, 1994; Sperber, 1996; see also Richerson & Boyd, 2005; Shennan, 2002, for different approaches to these issues). In addition, more than simply mnemic advantage may be at work. Thus, Tweney, Upal, Gonce, Slone, and Edwards (2007, in press) have shown that counterintuitive concepts can be used in creative writing tasks in ways that supplement Bartlett’s account of the role of remembering in social and cultural processes.

Franks (2003) has argued that the counterintuitive nature of religious concepts requires more than knowledge about the specific content of those concepts. Thus, in addition to their oddity, religious concepts are connected to other beliefs and to propositional attitudes (hopes, fears, and so on). Franks thus expressed concern that the cognitive-anthropological approach to explaining religious concepts may be insufficient if restricted to the cognitive aspects of the specific concepts themselves. Similarly, Whitehouse (2004) has argued for the “specialness“ of religious concepts, in that they require a network of costly social “mnemonic supports.” The examples provided by Franks and Whitehouse in support of their arguments are all derived from established, “culturally filtered,” beliefs—the exact relation of such beliefs to the online cognition that we have investigated remains an open question. Even so, our results suggest the possibility of a broader and more comprehensive cognitive account of established religious beliefs, one that can incorporate links to propositional attitudes and other beliefs. In terms of our theory, such relationships can be studied empirically by establishing their role in affecting the predictability and postdictability of religious counterintuitive concepts in appropriate contexts.

Most recent work has concentrated on showing that minimally counterintuitive concepts are indeed widespread (Barrett, 1997; Lisdorf, 2004) and that they are more memorable (Barrett & Nyhoff, 2001; Boyer & Ramble, 2001; Atran & Norenzayan, 2005). The focus of this work has been on the semantic and conceptual properties of the concepts themselves and less attention has been paid to the details of the cognitive processes involved and hence the context in which such processes take place. The current paper thus adds a crucial piece to this puzzle by presenting a model that explains the benefits of a memory capacity that leads to better recall for counterintuitive concepts and by showing how such benefits emerge in the real-time process of comprehending a narrative. Thus, like Bartlett’s, our account seeks to understand the dynamic processes of remembering in a way which will permit integration with cultural concepts and with models of cultural change.10 Most importantly, our model confirms that memorability is not an inherent property of a concept, rather it is a property of the concept, the context in which the concept is presented, and the background knowledge that the comprehendor possesses about the concept. By confirming specific prediccations derived from these notions, our model extends and renders precise the intuitions of Bartlett and others.

Notes

1. We follow existing practice in using the term “concept” to refer to both commonly used terms, such as “cow,” and newly constructed combinations such as “flying cow.”
It could be argued that the new combinations should be regarded as “associations,” although this would not alter our present argument.

2. We use the term agent in the sense in which it is used in artificial intelligence literature (Russell & Norvig, 1995) as covering the entire space of intelligent autonomous entities that includes humans, human-designed intelligent systems such as computer programs and robots, and intelligent aliens.

3. Note that memorizing, even in case of computer systems such artificial intelligence knowledge-base systems or data-bases, does not mean simply writing things onto a blank slate. Adding new information to a knowledge-base involves complex process of determining what parts of the knowledge-base to modify and delete (Alchourron et al., 1985). Our main point here is to argue that there are two broad strategies available to adaptive problem solving agents when confronted with the problem of making accurate predictions about their environment.

4. Schank’s (1981) failure-driven memory model is a more detailed computational model than our approach. Several case-based computer systems inspired by Schank have been designed over the years (Kolodner, 1993). Minimally counterintuitive concepts in such systems are easier to recall because they are indexed more distinctly. However, maximally counterintuitive concepts, by causing even more expectation failures, would be indexed and retrieved even faster than minimally counterintuitive concepts by failure-driven case-based systems.

5. Various theories differ on the mechanisms of generating such explanations and on the extent of inferences that are generated online (Graesser, 1997).

6. This is not true in all situations, of course. In scientific thinking, particularly in exploratory research, the need for abductive generation of possible explanations can be challenging; see Magnani (2001) and Tweney (2006).

7. The minimal inference models (McKoon & Ratcliff, 1992) assume that only local coherence is automatically attempted while constructionist models (Graesser, Singer, & Trabasso, 1994) hypothesize that readers attempt to derive both local and global coherence while reading texts.

8. Similar to our account, Franks (2003) suggests that counterintuitive concepts enjoy memory advantages because of the multiple possible interpretations that such conceptual combinations engender. However, unlike our comprehension-based account where postdiction process attempts to resolve contradictions among the properties of constituent concepts, Franks argues that many religious conceptual combinations are cognitively attractive because they leave such contradictions unresolved. He presents a number of examples of Catholic concepts which are treated as religious mysteries which are never explicitly resolved. However, Franks does admit the possibility that these “mysteries” may be resolved by different people differently in different contexts as do the readers of narratives containing counterintuitive concepts. Dennett (2006) suggests that "mysteries" in religious belief systems serve to stabilize counterintuitive concepts to prevent change over time.

9. For the present studies, we assume that this is constant across items. Actually computing the likelihood of C being encountered in the near future would require attention to a variety of factors that could be expected to vary across situations, including the
base rate of particular expectations, and the way these are modified by incoming information.

10. Tweney, Upal, Gonce, Slone, and Edwards (2007, in press) have shown that counterintuitive concepts can be used in creative writing tasks in ways that supplement Bartlett’s account of the role of remembering in social and cultural processes.

Acknowledgment

The authors would like to acknowledge the contributions of other members of the “I-75 Culture & Cognition Group” and Grounds For Thought Café in Bowling Green, OH for providing the stimulating atmosphere for discussion of the issues addressed in this paper. In particular, we’d like to thank Kristin Edwards for coding the results. Bradley Franks and two anonymous reviewers provided helpful comments that have greatly strengthened the paper.

References


Appendix A: Part of a questionnaire used in Experiment I

Topic 1: Singing dancing telephone

What can the phrase, “singing dancing telephone,” possibly refer to?
On a scale of 1 to 10 please indicate how easy or hard it was for you to answer the above question by circling the appropriate number.

<table>
<thead>
<tr>
<th>Very Easy</th>
<th>Little Harder</th>
<th>Much Harder</th>
<th>Impossible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Topic 2: Sobbing melting seaweeds

The sobbing melting seaweeds of Polynesia were recently discovered by the oceanographers. By melting parts of their stems into the blue ocean water, they produce a sobbing sound as ocean waves pass them.
Please turn the page to answer questions about this topic.
What can the phrase, “sobbing melting seaweeds,” possibly refer to?
On a scale of 1 to 10 please indicate how easy or hard it was for you to answer the above question by circling the appropriate number.

<table>
<thead>
<tr>
<th>Very Easy</th>
<th>Little Harder</th>
<th>Much Harder</th>
<th>Impossible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Topic 3: Flowering moving car

What can the phrase, “flowering moving car,” possibly refer to?
On a scale of 1 to 10 please indicate how easy or hard it was for you to answer the above question by circling the appropriate number.

<table>
<thead>
<tr>
<th>Very Easy</th>
<th>Little Harder</th>
<th>Much Harder</th>
<th>Impossible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Appendix B: Counterintuitive-supportive version of the story used in Experiment 2

The Adventures of Mr. Wurg

Mr. Wurg was about to be sent as an ambassador to the Razon 5 Galaxy. He wanted to know what things are like over there before leaving. So he went to the Arts and Science Museum to see the collection of artifacts you can find in houses on Razon 5. Razonians love smart object technology. Most Razonian objects have digital sensors and artificial intelligence chips
embedded in them to allow them to perceive their environment and act intelligently. Some older Razionian objects are dumb like the objects you find in most earthly homes.

Razonians have objects that are manufactured in small workshops. There are small places where Razonians make them.

They have robots that can cook food when asked to do so.

They have objects that can be taken apart to be fixed. You can take a part out to change it.

They have objects fitted with special sensors that can see everything that is going on around them.

They have objects made with parts from other objects. Someone took bits of other objects to make them.

They have objects that use their wheels to move back into their appointed place in the storage if they are left unused for a certain period of time.

They have objects designed by engineers. These people made a blueprint of the objects before making them.

They have objects that sense their owners’ biometric imprint and greet them as soon as they enter the dwelling.

They have objects that run away when strangers (i.e., people with different biometric imprints than their owners) try to grab them.

There are objects made by people as a hobby.

They have objects that use their sensors to detect that they are inside the house and start making high pitched alarm sounds if someone tries to remove them from their dwellings.

They have objects made of a special metal that’s hard to melt. It is difficult to make objects in that metal.

Mr. Wurg left the building and went to his office to ponder all of the things that can be found on Razon 5.

Appendix C: Modified (counterintuitive-supportive) version of Barrett & Nyhof’s (2001) “The Journey Home” Story

My Dream

I have always been fascinated by dreams. I have always wondered as to why some of our dreams are so different from our everyday experience; why are the laws of nature are violated so often in our dreams even though that never happens when we are awake? I remember that in some of my childhood dreams, our puppy Jack would to talk to me. In others, I would fly through the air. However, this particular dream must have been more bizarre than others because I wrote it down in my diary when I woke up.

July 7, 1985. Last night, I dreamed that my sister and I were walking home from school. We saw a cat belonging to one of our friends. The cat crouched on the front lawn as she composed a symphony. Since the cat was completely absorbed in her work, we continued on our way, chatting about what we had learned in school that day, until a beautiful rose jumped right in front of us. We knew that we must get home soon before our mother starts worrying so we slipped away from the rose.
We had hardly moved when we noticed an earthworm crossing the sidewalk. I picked it up and felt its slimy texture and the squirm of its movement. I put it on the grass on the other side of the pavement and we continued on our way. As we were walking, my sister’s shoes sprouted roots which broke up the pavement below and impaired her movement. I remembered that she had had this trouble with these shoes before and carried a pocket knife in order to cut the roots. While she was uprooting herself, a crumpled piece of newspaper blew past me in the cool breeze, almost brushing my leg.

The leaves, which had fallen from the trees a while ago, were brittle and crackled under our feet. We swished our way through the numerous leaves. I grabbed a bunch of leaves and throw them at my sister. I was getting kind of hungry so I reached into my backpack to retrieve a bag of carrots which I had not eaten during lunch. I was just about to take a bite of one of the carrots when it screamed, “Stop!” I decided to spare the carrot and put it back in the bag.

We continued on our way until my sister paused to notice a bright red bird perched on the branch of a nearby tree. She said that the red feathers of the bird signified that it was male. I moved to catch the bird but it flew away. My sister tried chasing it. She crossed the road and started running along a rolling field following the bird.

My sister was getting farther away from me. The sky had become darker, a few snow flakes had started falling, and the wind had picked up quite a bit. I was getting really cold and worried. That is when I saw Jack. He asked me if I needed help getting home and offered me a ride. I jumped on his back. He jumped into the air and started flying. My sister grabbed my right leg as I flew by her. I almost fell off the dog but I quickly grabbed Jack’s ears. Soon we caught sight of the white picket fence in the front yard. We flew past the fence and landed on our veranda. That’s when my mom woke me up and told me to get ready for school.