Melting Lizards and Crying Mailboxes: Children's Preferential Recall of Minimally Counterintuitive Concepts

Running head: RECALL OF MINIMALLY COUNTERINTUITIVE CONCEPTS

Melting lizards and crying mailboxes: Children's preferential recall of minimally counterintuitive concepts

Konika Banerjee*, Omar S. Haque, & Elizabeth S. Spelke

Harvard University, Department of Psychology

Keywords: Counterintuitive; Concepts; Religion; Evolution; Culture; Memory

* Corresponding Author.

Konika Banerjee is now at Yale University, Department of Psychology.

Tel.: 203 436 1547

2 Hillhouse Avenue, New Haven CT 06520, USA

E-mail addresses: konika.banerjee@yale.edu (K. Banerjee), haque@wjh.harvard.edu (O.S. Haque), spelke@wjh.harvard.edu (E.S. Spelke).

Abstract

Previous research with adults suggests that a catalog of minimally counterintuitive concepts, which underlies supernatural or religious concepts, may constitute a cognitive optimum, and is therefore cognitively encoded and culturally transmitted more successfully than either entirely intuitive concepts or maximally counterintuitive concepts. The current study examines whether children's concept recall similarly is sensitive to the degree of conceptual counterintuitiveness (operationalized as a concept's number of ontological domain violations) for items presented in the context of a fictional narrative. Seven-to-nine-year old children who listened to a story including both intuitive and counterintuitive concepts recalled the counterintuitive concepts containing one (Experiment 1) or two (Experiment 2) but not three (Experiment 3) violations of intuitive ontological expectations significantly more and in greater detail than the intuitive concepts, both immediately after hearing the story and one week later. We conclude that one or two violations of expectation may be a cognitive optimum for children: they are more inferentially rich and therefore more memorable, whereas three or more violations diminish memorability for target concepts. These results suggest that the cognitive bias for minimally counterintuitive ideas is present and active early in human development, near the start of formal religious instruction. This finding supports a growing literature suggesting that diverse, early-emerging, evolved psychological biases predispose humans to hold and perform religious beliefs and practices whose primary form and content is not derived from arbitrary custom or the social environment alone.

1. Introduction

1.1. Supernatural concepts and minimally counterintuitive ideas

Accumulating evidence from cognitive psychology suggests that religious beliefs are a by-product of a suite of ordinary, early-emerging cognitive mechanisms that originally evolved for non-religious purposes (Atran & Norenzayan, 2004; Barrett, 2000, 2004; Bloom, 2007; Boyer, 1994, 2000, 2003; Pyysiäinen & Hauser, 2010). For instance, our ability to reason in complex but systematic ways about social partners, agency, animacy, and causation may be essential cognitive foundations for inferences about the existence, intentions and everyday activities of supernatural deities, spirits, souls, ghosts, and ancestors. As a result, religious concepts may be predictable elaborations on the operations of natural human intuitive ontological systems (Boyer & H.C. Barrett, 2005; Spelke & Kinzler, 2007; Pinker, 2007), recruited for supporting supernatural beliefs. The forms and contents of religious concepts are therefore likely to be informed by these systems, and subject to the same types of cognitive constraints that apply to non-religious concepts (Atran & Norenzayan, 2004; Barrett, 2000, 2004; Boyer, 1994, 2000, 2003; Boyer & Ramble, 2001; Sperber & Hirschfeld, 2004).

One productive place to search for culturally successful religious templates is among supernatural concepts. Supernatural concepts form a major component of most religious traditions, texts, and beliefs, and they share a number of recurrent characteristics in every culture in which they have been studied (Atran, 2002; Barrett & Nyhof, 2001; Boyer, 1994, 2000, 2001; Boyer & Ramble, 2001; Brown, 1991; Pyysiäinen, Lindeman, & Honkela, 2003). Boyer (1994, 2000, 2001) has explained both the prevalence and characteristics of supernatural concepts by appealing to the underlying cognitive architectures that create them. Supernatural concepts, he

suggests, are "minimally counterintuitive": they violate one or a few of our psychological, biological, and physical intuitive ontological expectations about the properties of persons, non-human animals, plants, artifacts, or natural objects, while conforming to all other default ontological assumptions.

On this account, our penchant for the supernatural depends on our intuitive, default assumptions about the natural world. Default folk ontological assumptions are automatically generated, near-universal, early-emerging, mostly unconscious, domain-level inferences about the unique properties of persons, non-human animals, plants, artifacts, and natural objects (Boyer & H.C. Barrett, 2005). For example, as a part of folk physics, humans automatically assume that no two solid objects can occupy the same physical location and, as part of folk biology, that all living creatures must eventually die. These foundational ontological theories are built up from core knowledge systems that support representations of objects, actions, number, space, and social partners (Spelke & Kinzler, 2007; Pinker, 2007). As an example of a minimally counterintuitive concept, a ghost is a person (with all the default expectations about folk psychology), but one that can walk through walls (violating one aspect of folk physics). As a second example, a virgin mother is a regular mother who violates one aspect of folk biology because she gives birth to a child without prior sexual contact with any man.

1.2. Minimally counterintuitive ideas and cultural transmission

Previous work with adults from a number of different cultures in North America, Europe, Africa and Asia suggests that the catalog of minimally counterintuitive concepts, which underlies supernatural or religious concepts, may be especially memorable (e.g., Barrett & Nyhof, 2001;

Boyer & Ramble, 2001). As such, minimally counterintuitive concepts enjoy a privileged representational status, within and between cultures, over multiple iterations of cultural transmission. However, what is it about minimally counterintuitive concepts that makes them so memorable and so easily transmissible by humans with evolved minds like ours?

According to Boyer (1994, 2000, 2001), minimally counterintuitive concepts meet a kind of cognitive optimum that is a trade-off in the use of our attention and our memory capacities in representing ideas. Specifically, minimally counterintuitive concepts (a virgin mother) demand more attention than entirely intuitive concepts (a non-virgin mother) because they differ slightly from the types of conceptual templates and schemas that we already possess; i.e., they are schema-inconsistent. As a result, they cannot be automatically assimilated with pre-existing schematic knowledge and instead must be processed and encoded in greater detail than entirely intuitive, or schema-consistent, concepts. However, because minimally counterintuitive concepts only violate our intuitive expectations in a narrow and constrained way, we are still able to draw meaningful inferences about how the objects or agents in question are likely to behave (Boyer, 1994, 2000, 2001; Boyer & Ramble, 2001; Barrett & Nyhof, 2001). In contrast, a maximally counterintuitive concept (for example, a virgin, liquid mother who never engages in goal-directed thought) would be attention demanding, but too taxing to short-term memory and therefore would be harder to remember than a minimally counterintuitive concept. Thus, minimally counterintuitive concepts meet a cognitive optimum, because a *limited* number of counterintuitive violations of intuitive ontological expectations—as opposed to no violations, or many violations—makes a representation more memorable. As a result, minimally counterintuitive concepts are retained, reproduced, and transmitted more successfully than entirely intuitive concepts or maximally counterintuitive ones.

One might argue that there is something inherently religious about supernatural concepts that *a priori* makes them more memorable. From the perspective of an epidemiology of representations (Sperber, 1985, 1996; Boyer, 1994), however, such a claim may get things backwards. Rather, there is a certain class of concepts that happens to trade off attention and memory in a peculiar but inferentially rich manner. These concepts are more easily transmitted, and so over time we have come to weave these supernatural concepts into our culture-specific systems of belief. This approach helps explain why minimally counterintuitive concepts (especially of agents with beliefs, desires and/or intentions) are disproportionately represented not only in religious texts, but also in successfully transmitted literatures of all kinds—including folklores, mythologies, fairy-tales, novels, songs and poetry—whether religious or secular.

1.3. Additional past work on minimally counterintuitive concepts

The current experimental literature offers a nuanced picture of the factors that may influence preferential recall of minimally counterintuitive over intuitive concepts. In studies of adults, participants typically are asked to read or listen to a list of phrases or a narrative containing minimally counterintuitive, intuitive, and maximally counterintuitive target items. After completing a short, unrelated distractor task, participants must recall as many phrases from the original list as possible, or retell the original story while preserving as much detail as possible. Experimenters then measure the percentage of total minimally counterintuitive versus intuitive versus maximally counterintuitive concepts successfully recalled. Among the relevant reported variables affecting differential recall are: whether items are presented in context-free lists or context-rich narratives (Upal, Gonce, Tweney, & Slone, 2007; Gonce, Upal, Slone, & Tweney,

2006), the method of item transmission (orally via social partners or in written text format) (Barrett & Nyhof, 2001), and whether the recall task is administered immediately after exposure to the target concepts or after a longer imposed delay (Barrett & Nyhof, 2001; Norenzayan, Atran, Faulkner, & Schaller, 2006).

Specifically, recent experiments suggest that the memorability of minimally counterintuitive concepts depends not only on the concepts themselves, but also on the context in which the concepts are presented, and on the existing schematic knowledge that is activated by that context (Upal et al., 2007; Gonce et al., 2006). As a result, though minimally counterintuitive concepts presented within a narrative context are consistently recalled more reliably and accurately than intuitive concepts (Barrett & Nyhof, 2001; Boyer & Ramble, 2001), this memory advantage disappears when these same concepts are instead presented in context-free lists that prevent activation of background schematic knowledge (Gonce et al., 2006; Upal et al., 2007). Because context provides valuable information about a narrative's discourse genre, the lack of context may alter people's expectations about target concepts, and subsequently, their memory for those concepts (Gonce et al., 2006; Zwaan, 1993). Consequently, in emphasizing the role of context as well as concepts, Gonce et al. (2006) have suggested that the appropriate unit of study in conceptual recall tasks is a narrative rather than a list of items alone.

There is also evidence that recall of minimally counterintuitive concepts with associated context is significantly better than recall of intuitive concepts as the delay between encoding and recall *increases*. Compared to intuitive concepts, minimally counterintuitive concepts are recalled better after an imposed delay of weeks or months, and thus the recall advantage for minimally counterintuitive concepts is *larger* on delayed than on immediate recall tasks (Gonce et al., 2006). That minimally counterintuitive concepts are recalled better than intuitive concepts after a long

delay suggests a potential "differential fitness" of these concepts, whereby memory for intuitive concepts decays more rapidly over time than does memory for minimally counterintuitive concepts (Gonce et al., 2006). This finding of differential fitness may account, in part, for the long-term success of both religious and non-religious cultural narratives and concepts that conform to the minimally counterintuitive template.

1.4. Hypothesis and predictions

Past experimental work on these topics has been entirely with adults, and no research to date has examined the differential recall of minimally counterintuitive, intuitive, and maximally counterintuitive concepts in children. The main aim of this investigation is to ask: Do children, like adults, also demonstrate a relationship between the recall of concepts and the number of ontological domain violations of items recalled in the context of a fictional narrative? If such a relationship exists in children, then how many violations are required to show a boost in memorability, and how many further violations are required to undo this effect?

The prediction that children, like adults, will find moderate departures from intuitive concepts more memorable gains plausibility from findings that the foundational folk psychological, biological and physical theories that inform our intuitive expectations about the natural world emerge early in life, with some even present in newborns (Boyer & H.C. Barrett, 2005; Spelke & Kinzler, 2007). As a result, it is reasonable to hypothesize that children, like adults, are capable of recognizing violations of categorical knowledge, and so may also show a heightened sensitivity for recalling concepts at, but not beyond or before, a particular level of counterintuitiveness (as measured as number of ontological domain violations of items recalled in

the context of a fictional narrative). Such a finding would be consistent with the widespread nature of minimally counterintuitive-like concepts in children's fairy tales and folk stories (e.g. Santa Claus, the Tooth Fairy) (Grimm & Grimm, 1869/2003). It is possible that because these concepts are highly salient and memorable, children's stories—including religious ones—and folktales that contain them are especially likely to be culturally successful (Norenzayan et al., 2006). Evidence for a privileged status for counterintuitive concepts in children's memory would speak to an important developmental continuity in specific cognitive biases that account for supernatural beliefs.

Alternatively, children may not distinguish between concepts based on their number of violations of intuitive ontological expectations, and they may fail to demonstrate superior memory for minimally counterintuitive concepts for several reasons. First, children know vastly less about the natural world than adults, and their primary cognitive task, arguably, is to gain such knowledge. For this reason, children may attend more to intuitive concepts than do adults. Second, global memory and recall demands may be greater for children than for adults, rendering differences between recall of concepts with different numbers of violations of intuitive ontological expectations undetectable in children until later in development. The present experiments focus, therefore, on children of early school age (7 to 9 years). Such children have already built elaborate systems of intuitive knowledge about the natural world (Wellman & Gelman, 1992), a critical prerequisite for testing whether children are sensitive to counterintuitive violations derived from intuitive theories of biology, psychology, and physics. In addition, we selected 7-9-year old children because children of this age show considerable competence on sophisticated memory tasks (Beardsworth & Bishop, 1994; Gambrell & Jawitz, 1993; Courage & Cowan, 2008). Although even younger children may theoretically also be sensitive to conceptual

counterintuitiveness, the task demands of the current experiments are better suited to the memory and story comprehensions capacities of older children. For instance, we sought to avoid potential recall floor effects in delayed story recall amongst younger children, which may have prevented analysis of differential decay of intuitive and counterintuitive concepts over the course of one week. We therefore predict that 7-9-year old children will show superior recall of counterintuitive concepts from narratives.

1.5. The present experiments

In three experiments, we ask whether children exhibit superior recall of counterintuitive concepts, compared to entirely intuitive concepts, when these concepts are presented within the narrative context of a fictional story. In addition, we ask whether, for children, a relationship exists between degree of conceptual counterintuitiveness and concept recall, such that concepts containing a limited number of violations of intuitive ontology, but not more, are recalled better than intuitive concepts. We explore the effect of an imposed time delay on concept recall by administering two separate recall tasks, the first task immediately after hearing the story, and the second task one week later.

In Experiment 1, we present children with a narrative containing both counterintuitive concepts that entail a single violation of intuitive ontological expectations and intuitive concepts containing no such violation. In Experiments 2 and 3, we test whether children's memory for counterintuitive concepts varies as a function of the number of violations described by adding a second (Exp. 2) and then a third (Exp. 3) violation of intuitive ontological expectations to the target counterintuitive concepts. Taken together, Experiments 1, 2, and 3 allow us to probe

whether a cognitive threshold in children favors minimally counterintuitive concepts over more counterintuitive ones. Following Boyer's model for adults (1994, 2000), we predict that concepts that contain a limited number of counterintuitive violations achieve an optimal trade off between attention and inferential productivity, and should therefore be recalled significantly better than intuitive concepts by children. However, we expect that concepts with many counterintuitive violations will inhibit inference generation at the expense of concept recall and will be cognitively encoded more poorly than minimally counterintuitive ones. We therefore predict that when compared to intuitive concepts, maximally counterintuitive concepts should not demonstrate the recall advantage characteristic of minimally counterintuitive concepts.

2. Experiment 1: Children's memory for intuitive concepts vs. conceptions with one counterintuitive violation.

2.1. Participants

Participants were 23 boys and 19 girls, ranging in age from 7 to 9 years of age (M = 8.16 years, SD =.83, range = 7.03-9.67). All children were English-speakers, though 19% were raised in multilingual homes. We collected information from parents about religious affiliation and religiosity for 57.1% of children (N = 24) who participated in Experiment 1. These children were 54.2% Christian, 12.5% Jewish, 12.5% Muslim, 4.2% Other, and 16.7% Atheist/Agnostic. Parents also rated how important it was to them to raise their children in a religious tradition, using a 1-7 scale, anchored at "not at all important" (1) and "extremely important" (7). Ratings were bimodally distributed, with most parents reporting either a 2 (25%) or a 6 (20.8%),

suggesting that participants tended to come from families in which being raised in a religious tradition was either fairly important or else fairly unimportant. In addition, parents indicated how often their child attended services at a place of worship by selecting one of the following options: every week or more often, once or twice a month, a few times a year, once a year, less than once a year, never, and not a member of a religion. The modal response was one or twice a month. Children participated in a single session in the lab for the first part of the study (immediate recall task) and were called at their homes one week later for the second part (delayed recall task).

2.2. Materials

All the concepts were introduced to children in the context of a single, short fictional story. The story structure was loosely modeled after the narrative used by Barrett & Nyhof (2001, Experiment 3). We created two versions of the story. Both story versions had an identical plot in which two children encounter a number of objects (target items) while exploring a new neighborhood (i.e., the story was not set in an unfamiliar world in which counterintuitive concepts might be expected). The target items were basic-level objects (e.g., kinds of plants, animals, and artifacts), each associated with two descriptors (see below). Each story contained a total of twelve target items, consisting of six target nouns paired with an intuitive description, and six target nouns paired with a counterintuitive description. The intuitive and counterintuitive descriptions were counterbalanced across the two story versions so that target nouns paired with an intuitive description in version one were paired with a counterintuitive description in version two and vice versa. This ensured that the order of target items, which was the same across both story versions, would not influence recall patterns. We also included two versions of the story to

ensure that recall effects were not a result of the variable saliency or memorability of particular target nouns, but rather a function of target item descriptor type (intuitive or counterintuitive).

The six counterintuitive descriptions in each story included as their ontological domain violations three breaches and three transfers, as described by Boyer (2000). Specifically, breaches violate the assumed psychological, biological, or physical properties of entities within any ontological category. For example, "a desk that blips in and out of existence" describes a counterintuitive breach of physics (sporadic non-existence) that cannot be true of members of any ontological category. In contrast, transfers apply a property characteristic of members of one ontological category to members of a different ontological category for which that property is inappropriate. For instance, "a desk that feels happy" describes a counterintuitive transfer of a psychological property (the capacity for emotion) to an inanimate human-made artifact that typically lacks psychological processes.

Target items were created by first generating a list of random nouns belonging to four distinct domain categories: 1 animal, 2 plants, 1 natural object, and 2 human-made artifacts. Each target item was then paired with an intuitive description consisting of two adjectival clauses that both adhered to that item's default domain-level properties (e.g., a lizard that (1) ate insects off the ground and (2) crawled around quickly on all four of its feet). Each noun was also paired with a counterintuitive description consisting of one intuitive descriptive clause with no property violation and a second counterintuitive descriptive clause in which a single domain-level property was violated (e.g., a lizard that (1) had a long, thin tail and (2) could never die no matter how old it was). The total length (number of words) of the intuitive and counterintuitive descriptions was matched for each target item. The six counterintuitive target items, as a whole, included two violations each of psychological, biological, or physical ontological properties. Thus, we tested

whether violation type influenced the likelihood of recall among different counterintuitive target items. Violation types always appeared in the same order for both story versions.

To validate our own violation type designations, we asked 23 hypothesis-blind adults to provide independent ratings of the counterintuitive properties described in the counterintuitive items. Adults were told only to choose whether they believed each property described represented a violation of psychology, biology, or physics. We intentionally withheld more detailed instructions because we sought to elicit adults' spontaneous, intuitive judgments. Adult ratings of the counterintuitive property violation types (psychology, biology, or physics) were sometimes inconsistent, and a majority consensus was achieved for only 4 of the 6 counterintuitive violations. Notably, the two violations for which there was substantial disagreement were the properties we designated as violations of psychology. Adults nearly equally categorized these properties as violations of psychology or biology, perhaps reflecting their inclination to construe mental states attributed to inanimate objects (icicle/rock and stop sign/mailbox) as being rooted in biology.

Next, a second noun was paired with each of the original target nouns. Both paired nouns belonged to the same domain category, (e.g., both were plants or both were human-made artifacts) so that the associated intuitive and counterintuitive descriptions would be equally applicable to both nouns. In story version one, the first noun in the pair was matched with the intuitive description and the second noun was matched with the counterintuitive description. This was reversed in story version two. For both story versions, each target noun was mentioned first in an introductory sentence that tied it to the story plot but had no associated description, and then a second time in an elaborative sentence that included either a counterintuitive or intuitive description. The target items and descriptions used in both story versions are presented in Table

1.

----------Insert Table 1 about here ----------

2.3. Design and procedures

2.3.1. Story task

Children were assigned to hear one of two versions of the story. Assignments were controlled for gender, so that an approximately equal number of boys and girls heard each story version. Children were brought into the experimental room and were seated across from the experimenter. Parents did not typically accompany the children into the room, although four parents specifically requested being present to observe. The experimenter told the children that she would read them a short story out loud and asked them to listen to the story carefully and to try to imagine the events described in the story while listening. Oral transmission of the story was desirable both because it most closely mimics the way in which most cultural narratives have historically been transmitted to children (Goody, 1987; Graham, 1993), and also because it avoided potential confounds relating to variable reading comprehension and speed among 7-9-year olds. Children were told that they would be asked questions about the story later on in the study. The experimenter then read the roughly 600-word story out loud for approximately four minutes. See Appendix A for full texts of both story versions for Experiment 1. The same experimenter ran all participants in Experiment 1.

2.3.2. Distractor task

Immediately after hearing the story, children completed a short, unrelated computer task intended to temporarily distract them from the story prior to the recall task. Children saw three

angles on a computer screen and were asked to pick which two of the three angles shown looked the most similar to each other. The task consisted of a total of 24 trials, including 12 two-dimensional angle trials and 12 three-dimensional angle trials. The two-dimensional angles were formed by two lines, and the three-dimensional angles were formed by two walls of a room. For each trial, one angle was presented at the top of the computer screen and two other angles were presented side by side directly underneath. Children pressed one of two buttons on the keyboard to indicate whether they thought the picture on the left or right-hand side of the screen looked the most like the picture at the top of the screen. The similarity distractor task purposely avoided working memory demands so as to prevent potential interference with the story recall task.

2.3.3. Story recall task

After completing the distractor task, children were instructed to think back to the story they heard earlier and to try to remember as many details from it as possible. The experimenter asked children nine prompting questions that required them to recall plot events from specific parts of the story (Table 2). Most questions were designed to elicit recall of between one and three of the target items. Questions 2 through 7 were each asked once, and then repeated again if the child did not recall all of the target items primed by the question. This methodology was used because children's overall recall of story details was much better when they were asked the prompting questions more than once. The experimenter asked questions 1 and 8 only once because these questions were not intended to prime recall of any particular target item, but rather were included to encourage children to recall the story's plot more generally. Question 9, which probed overall story recall, was asked twice only if children provided a response the first time the question was asked. Plot-relevant prompting questions were used rather than a more general instruction to

recall the story in its entirety because an earlier pilot study confirmed that specific plot-relevant prompting questions elicited far superior overall recall.

If children successfully recalled a target noun, they were asked what else they could remember about that noun. The purpose of this elaboration prompt was to elicit greater recall of the descriptive clauses associated with each target noun. A prior pilot study found that children tended to omit descriptive clauses when recalling target items unless specifically prompted to elaborate on what they could remember *about* successfully recalled target nouns. In general, the experimenter gave children verbal positive reinforcement whenever they recalled story details, and did so regardless of recall accuracy, in order to uniformly encourage further recall of the story. A transcript of children's responses was recorded live by the experimenter, and the entire lab session was video recorded.

----------Insert Table 2 about here ----------

2.3.4. Delayed recall task

During the initial lab visit, the experimenter told parents that she would call their children at home one week after their lab session to complete the delayed story recall task. Parents were explicitly told not to tell their children that they would be contacted again. During the follow-up phone call, the experimenter asked the children if they remembered the story they heard in the lab one week earlier. Children were then asked to think back to the story and to try to remember as many details from it as possible. The first portion of the delayed recall task was strictly free recall in which the children were asked to list everything they could remember about the story without the experimenter's use of any prompting questions. Following the free recall portion of the task, the experimenter asked children the same nine prompting questions used in the immediate recall task. Once again, the experimenter asked each question twice if the child

failed to recall each of the target items primed by each question after the first time it was asked. Children were also again asked to elaborate on any target nouns that they successfully recalled. Again, the experimenter gave children verbal positive reinforcement whenever they recalled story details, and regardless of accuracy, in order to uniformly encourage further recall of the story.

2.4. Results

2.4.1. Recall coding

The primary experimenter and a hypothesis-blind research assistant independently recorded transcripts of each child's responses on the immediate recall task and coded them for target item recall. The research assistant produced transcripts by viewing video recordings of 93% of the lab sessions. For the three lab sessions (7%) that were missing video recordings, the research assistant used the primary experimenter's live-recorded transcripts to code children's responses. Because the delayed recall task was conducted by phone, the research assistant coded children's delayed recall using the primary experimenter's live-recorded transcripts of the phone calls.

Target item recall was coded in two ways. First, each target item was assigned a total of three possible recall points: one for recalling the target noun, and one for recalling each of the two descriptive clauses. Children were given a recall score between 0 and 3 for each of the twelve target items. If a child recalled the general idea of a target noun or descriptive clause—by either situating the noun correctly within the story's plot or by correctly identifying the associated descriptive clauses—the item was coded as having been accurately recalled. In order for a descriptive clause to be coded as accurately recalled, it needed to preserve the relevant

counterintuitive or intuitive featural cues. In other words, accurately recalled counterintuitive clauses had to include a violation of categorical expectation, but accurately recalled intuitive clauses could not include this type of violation and had to include their own intuitive descriptor that did not violate categorical expectations. Inter-coder reliability was very high, (99.1% agreement on all recall points awarded), and so the primary experimenter's judgments were used in the minority of cases in which there was disagreement.

Second, because children tended to recall the target item nouns more reliably than the descriptive clauses, we separately coded only the total number of nouns recalled that were associated with counterintuitive descriptions or intuitive descriptions (hereafter referred to as counterintuitive nouns and intuitive nouns). Children were given a score of one for each of the twelve target nouns recalled. We were interested in whether children recalled more of the counterintuitive nouns than the intuitive nouns even if they failed to recall the nouns' associated descriptive clauses.

2.4.2. General results

First, we performed preliminary analyses to test for potential effects of several variables that were not predicted to have a systematic effect on childrens' target item recall. These variables included story version, gender, religiosity, age, time required to complete the distractor task, and the experimenter's reading of the target items. These analyses were conducted both to determine whether participant demographics influenced target item recall and also to validate our experimental methodology.

We randomly assigned children to hear one of two versions of the fictional story in order to counterbalance each target noun's pairing with a counterintuitive and an intuitive descriptor

across versions. This design allowed us to test whether target item recall was potentially influenced by the variable saliency of the particular target nouns that we selected, rather than by target item descriptor type (intuitive or counterintuitive). As expected, repeated measures ANOVAs revealed no effect of story version on overall target item recall (nouns + descriptive clauses) either immediately, $F(1, 40) = .46$, $p = .50$, or after one week, $F(1, 40) = .63$, $p = .43$. Similarly, we found no significant interaction between story version and target item type (counterintuitive or intuitive) on recall either immediately, $F(1, 40) = .37$, $p = .54$, or after one week, $F(1, 40) = .14$, $p = .72$. Thus, recall effects were not influenced by the variable saliency of particular target nouns, and we collapsed across story versions for all subsequent analyses.

Next, we performed several analyses to determine whether recall effects may have been driven by differences in the childrens' demographics or by theoretically unimportant features of the experimental design. There was only a marginal effect of gender on overall target item recall for the immediate recall task, $F(1, 40) = 3.40$, $p = .073$, whereby girls recalled target items slightly better than boys, and a non-significant effect of gender on the delayed recall task, $F(1, 40) = 2.72$, $p = .11$. Participants' age was uncorrelated with overall target item recall for both the immediate, $r = .063$, $p = .69$, and delayed recall tasks, $r = .11$, $p = .53$. Because we collected religiosity information for only 57.1% of children in Experiment 1, we combined participants for whom we had this information across Experiments 1, 2, and 3 ($N = 53$) to test for the effects of religiosity on target item recall. Our measures of religiosity allowed us to assess whether extent of exposure to counterintuitive religious or supernatural concepts outside of the lab was related to performance on the target item recall tasks. Both overall counterintuitive and intuitive target item recall were uncorrelated with parental reports of importance of raising their child in a religious tradition and with frequency of attendance at religious services (all ps >.30), The number of

minutes required to complete the distractor task during the initial lab session varied across children (M = 4.03 minutes, SD = .93, range = 2.51-6.56) but was not correlated with overall immediate target item recall, r = .037, p = .82. In sum, variables of story version, gender, age, religiosity, and time to complete distractor task had no consistent effect on target item recall for Experiment 1 nor for any of the subsequent experiments, and they are therefore not discussed further.

In order to ensure that counterintuitive and intuitive items were read by the experimenter with equal emphasis, a hypothesis-blind coder reviewed video-taped recordings of 20% of all participant lab sessions for which data was collected across Experiments 1, 2, and 3[1]. The coder rated the experimenter's reading of each of the 12 target items on the following measures: volume, pitch, expressiveness, and experimenter facial expression while the item was being read. Because the camcorder angle sometimes recorded the experimenter's face in profile, the coder was asked to code for facial expression to the best of her ability, though some trials prevented the coder from observing a full frontal view of the experimenter's face. A 2x2 repeated measures ANOVA revealed that there were no differences in overall immediate recall of counterintuitive and intuitive target items between the lab sessions selected for coding and the uncoded sessions, $F(1, 101) = .49$, $p = .49$, confirming that a representative sample was selected. Repeated

---

[1] In addition, four of the coded lab sessions were from the control experiment described at the end of the results section for Experiment 1. A total of 21 participant lab sessions were coded. Because Experiments 1, 2, and 3 were run by a single experimenter and the control experiment was run by a different experimenter, experimenter was included as a between-subjects variable in subsequent analyses of target item reading.

measures ANOVAs revealed that there were no significant differences in how counterintuitive and intuitive items were read by the experimenter on any of the measures of interest, all *p*'s > .10.[2]

2.4.3. Immediate recall task

On average, children recalled 56.8% of the target item nouns (6.81 of 12 nouns) during the immediate story recall task. In terms of recall points (nouns + descriptive clauses), children scored an average of 14.1 out of a possible total of 36 points, with an average 7.29 points corresponding to recall of descriptive clauses. A repeated measures ANOVA revealed a significant main effect of target item type on noun recall, $F(1, 41) = 8.88$, $p = .005$, $\eta^2 = .178$. Children recalled nouns given a counterintuitive property significantly better than nouns given only intuitive properties. On average, children accurately recalled 64.7% of the counterintuitive nouns ($M = 3.88$ out of 6, $SD = 1.35$) but only 48.8% of intuitive nouns ($M = 2.93$ out of 6, $SD = 1.42$). We also found a significant effect of item type on overall target item recall points, $F(1,41) = 15.24$, $p < .001$, $\eta^2 = .27$. On average, children scored 8.40 out of a total of 18 possible recall points for counterintuitive items ($SD = 3.36$), but only 5.69 points for intuitive items ($SD = 3.02$). Overall, children recalled counterintuitive target items more frequently and in greater detail than intuitive target items. Across all subjects, children recalled target item nouns more often when paired with counterintuitive descriptions (64.7%, 163 out of 252 times) relative to when those same nouns were paired with intuitive descriptions (48.8%, 123 out of 252 times). In addition,

[2] There were also no experimenter by target item type interactions for any of these measures, all *p*'s > .12.

five of the top six most frequently recalled target item nouns were paired with counterintuitive descriptions.

Children frequently recalled target nouns without remembering any or all of those nouns' associated descriptions. There were no significant differences in the rate at which children recalled either zero or one adjectival clauses in conjunction with successfully recalled intuitive or counterintuitive target nouns. However, when children successfully recalled both adjectival clauses, they were more than twice as likely to do so in conjunction with counterintuitive nouns than with intuitive nouns, $F(1, 41) = 16.48$, $p < .001$, $\eta^2 = .29$. In other words, children were significantly better at recalling all of the associated detail provided in the narrative for counterintuitive concepts than for intuitive concepts. Additionally, when children recalled only one adjectival clause in association with a counterintuitive noun, they were more than twice as likely to recall the counterintuitive adjectival clause than the intuitive adjectival clause, $F(1, 41) = 7.55$, $p = .009$, $\eta^2 = .16$ (Fig. 1).

----------Insert Figure 1 about here ----------

Across all children, nouns paired with breaches were recalled more often (73.8% of all possible instances) than nouns paired with transfers (52.4% of all possible instances). To assess within-subjects differences in children's recall of breaches and transfers, each child was given a score between 0 and 3 for the number of transfers accurately recalled and a separate score between 0 and 3 for the number of breaches accurately recalled. A repeated measures ANOVA on these scores revealed that children recalled breaches significantly better than they recalled transfers, $F(1, 41) = 24.46$, $p < .001$, $\eta^2 = .37$ (Fig. 2). A repeated measures ANOVA also revealed a significant effect of counterintuitive violation type on recall, $F(2, 82) = 4.43$, $p = .015$, $\eta^2 = .10$. This effect was driven solely by a recall advantage for physics violations over

psychology violations, $F(1, 41) = 7.03$, $p = .011$, $\eta^2 = .15$.

----------Insert Figure 2 about here ----------

2.4.4. Delayed recall task

Of the 42 children who participated in the initial lab session, 36 were successfully contacted one week later to complete the delayed recall task. Within subjects, delayed recall was highly correlated with immediate recall, (nouns: $r = .77$, $p < .001$; points: $r = .81$, $p < .001$). However, children who completed both the immediate and delayed recall tasks recalled significantly more target nouns and recalled them in greater detail during the immediate recall task compared to the delayed recall task (nouns: $F(1, 35) = 11.17$, $p = .002$, $\eta^2 = .24$; points: $F(1, 35) = 42.67$, $p < .001$, $\eta^2 = .55$). Nevertheless, delayed target item recall remained relatively high in relation to immediate recall rates. On average, children who participated in both recall tasks recalled 6.61 of the possible 12 target item nouns during the initial lab visit, with recall dropping only slightly to 5.83 nouns one week later. These children scored an average of 13.67 out of 36 possible recall points during the immediate recall task, but only 10.55 points after one week.

One week after hearing the story, children recalled an average of 48.6% of the target nouns (5.83 of 12). There was a significant main effect of target item type on noun recall, $F(1, 35) = 4.58$, $p = .039$, $\eta^2 = .12$. Children accurately recalled 53.7% of the nouns with a counterintuitive property ($M = 3.22$ out of 6, $SD = 1.40$) but only 43.0% of the nouns with only intuitive properties ($M = 2.58$ out of 6, $SD = 1.40$). There was also a significant effect of item type on overall target item point recall, $F(1, 35) = 17.54$, $p < .001$, $\eta^2 = .33$. Children scored an average of 6.53 out of a total of 18 possible recall points for counterintuitive items ($SD = 3.11$), but only

4.03 points for intuitive items (*SD* = 2.71). Fig. 3 displays differences in overall counterintuitive versus intuitive target item recall points (nouns + descriptive clauses) for both the immediate and delayed recall tasks[3]. As in the immediate recall task, target nouns in the delayed recall task were recalled more often across children when paired with counterintuitive descriptions (53.7%, 116 out of 216 times) relative to when those same nouns were paired with intuitive descriptions (43.1%, 93 out of 216 times). In addition, four of the top five most frequently recalled target item nouns were paired with counterintuitive descriptions.

----------Insert Figure 3 about here ----------

We performed paired samples t-tests on the difference scores between children's recall for counterintuitive versus intuitive target items (nouns and points) to determine if there were differences in the magnitude of these effects between the immediate and delayed recall tasks. No differences in magnitude were observed for either noun recall, $t(35) = 1.43$, $p = .16$, nor recall points, $t(35) = -.50$, $p = .62$. Thus, there was no observed differential decay of counterintuitive and intuitive concept memory over the course of one week.

---

[3] This figure and all others that compare immediate target item recall to delayed target item recall include data from all children who completed each recall task in order to present all data collected. Because not all children who completed the immediate recall task also participated in the delayed recall task, more children are represented in the immediate recall task bars than in the delayed recall task bars. However, the relationships between immediate and delayed recall remain unchanged if only the children who completed both recall tasks are considered.

Paralleling the results of the immediate recall task, children recalled breaches significantly more frequently than they recalled transfers, $F(1, 35) = 6.70$, $p = .014$, $\eta^2 = .16$ (Fig. 2). In addition, there was a significant within-subjects effect of counterintuitive violation type (psychology, biology, or physics) on recall, $F(2, 70) = 4.62$, $p = .013$, $\eta^2 = .12$. This main effect was driven by a recall advantage for physics violations over psychology violations $F(1, 35) = 5.21$, $p = .029$, $\eta^2 = .13$ and for physics violations over biology violations, $F(1, 35) = 5.83$, $p = .021$, $\eta^2 = .14$.

2.4.5. Target item order control experiment

Although counterintuitive and intuitive target items were distributed throughout the first and second halves of the story, the first half of the story contained more counterintuitive target items than did the second half of the story. In order to determine if the obtained recall effect was driven instead by a primacy order effect (i.e., a recall advantage for items appearing earlier in the story), we ran twenty additional 7-9-year old children in a control experiment that used a slightly modified version of the story. In this version, counterintuitive and intuitive target item order was identical for both the first and second halves of the story. The story plot and target items were nearly identical to those used in the original story, with only minor modifications made to accommodate changes in target item position in order to maintain plot coherence. See Appendix B for the full text of the modified control story.

Inter-coder reliability between the experimenter and a hypothesis-blind coder on measures of children's recall was very high (99% agreement on all recall points awarded). Replicating the results of the immediate recall task in Experiment 1, children recalled the counterintuitive target items containing a single violation of intuitive ontological expectations significantly more often

(nouns: $F(1, 19) = 5.52$, $p = .03$, $\eta^2 = .23$) and in greater detail (recall points: $F(1,19) = 23.87$, $p < .001$, $\eta^2 = .56$) than the intuitive target items immediately after hearing the story.

In addition, we performed a multilevel ordered logistic regression to test for the effect of item type (intuitive versus counterintuitive) on children's recall points, controlling for target item order (1-12). The regression model specified children's recall point score on each of the twelve target items (0-3) as the dependent variable and target item type (intuitive versus counterintuitive) as the independent variable. The dependent variable had a multi-level structure with participant at the higher level and target item order (1-12) at the lower level. This model allowed for within-subject correlated responses at the lower level (target item order). The coefficient of target item type, which indicates the effect of target item type on children's mean recall scores, was negative, -.63, $SE = .24$, $p = .009$. Thus, the expected predicted recall score when the item type was intuitive (item type = 1) was significantly lower than when item type was counterintuitive (item type = 0). The results of this analysis further confirmed that children's expected recall scores were higher for counterintuitive items than for intuitive items, even after controlling for the target items' position in the story. Children's superior recall of counterintuitive concepts in Experiment 1 was therefore not an artifact of item order, as even when counterintuitive and intuitive items were more evenly distributed throughout the story and when item position in the story was controlled for statistically, a main effect of item type on recall still obtained.

2.5. Experiment 1 discussion

Children who listened to a fictional narrative containing both intuitive concepts and counterintuitive concepts containing a single violation of intuitive ontological expectations recalled the counterintuitive concepts more frequently, and in greater detail, both during an

immediate recall task and also one week later. Target nouns were recalled significantly better when paired with counterintuitive descriptions relative to when these same nouns were paired with intuitive descriptions. Thus, differences in recall could not be attributed to the varying saliency of the nouns themselves, but rather to their conceptual formulation within the story. These results suggest that for children, like adults, counterintuitive concepts containing a single violation of intuitive ontological expectations possess the recall advantage characteristic of minimally counterintuitive concepts. Thus, children also appear sensitive to the unique features of the minimally counterintuitive cognitive template that render the catalog of minimally counterintuitive concepts cognitively salient and memorable.

In addition, Experiment 1 offers initial evidence that among violations of ontological expectations, children recall counterintuitive breaches more reliably than counterintuitive transfers, suggesting that single breaches may be more cognitively salient than single transfers. Children also recalled violations of intuitive physics better than violations of intuitive psychology. These effects may be related, as both psychology violations were transfers and both physics violations were breaches. The two biology violations included one breach and one transfer, which may also account for why physics violations (both breaches) were recalled better than biology violations in the delayed recall task.

3. Experiment 2: Children's memory for intuitive concepts vs. conceptions with two counterintuitive violations.

Having established that concepts with one ontological violation were more memorable for children than concepts with no violations, we next probed children's memory for concepts with

two ontological violations. The method was the same as Experiment 1, except that each of the counterintuitive nouns was associated with two properties that violated basic ontological distinctions. Children's memory for these concepts, relative to entirely intuitive concepts, was compared as in Experiment 1.

3.1. Participants

Participants were 11 boys and 9 girls, ranging in age from 7-9 years of age (M = 8.63 years, SD = .94, range = 7.18-9.99). All children were English-speakers, though 50% were raised in multilingual homes. We collected information on religious affiliation and religiosity from parents for 80% of children (N = 16). These children were 50% Christian, 1.9% Muslim, 6.3% None, 6.3% Other, and 18.8% Atheist/Agnostic. Most children came from households in which parents either felt that it was extremely important or fairly important to raise their children in a religious tradition (47% reported a 6 or 7 on the 1-7 scale) or else that it was not at all important to do so (29.4% reported a 1 on the 1-7 scale). Most children (N = 10) attended religious services every week or more often or else once or twice a month. In all other ways, the participants in Experiment 2 were equivalent to those who participated in Experiment 1.

3.2. Materials

The fictional story and target concepts used in Experiment 2 were identical to those used in Experiment 1 with one exception. The single intuitive adjectival clause paired with each counterintuitive target item in Experiment 1 was replaced with a violation of intuitive ontology.

Accordingly, counterintuitive target nouns in Experiment 2 were paired with *two* counterintuitive violations rather than one. The second adjectival clause for all counterintuitive descriptions was matched in length (number of words) to the second adjectival clause of corresponding intuitive descriptions for all target nouns, except in one instance (icicle/rock) in which the clauses differed in length only by a single word. Intuitive target items remained unchanged.

Each counterintuitive item was designed to contain two different types of intuitive ontological violation (psychology, biology, or physics). For example, one violation of physics plus one violation of biology were permitted but two violations of biology were not. This constraint was intended to ensure that children encoded the two violations described for each counterintuitive item distinctly, and not simply as complementary elaborations on the same single violation. Furthermore, this approach allowed for a better test of whether concepts containing two non-overlapping violations of intuitive ontology meet the cognitive criteria of minimal, but not yet maximal, counterintuitiveness. Each counterintuitive item also contained one breach and one transfer[4].

---

[4] The definition of breach employed in the present studies differs from Boyer and Ramble's (2001) usage. We define breaches to be any psychological, biological, or physical feature that violates the assumed properties of members within any ontological category. Thus, under our definition, a "hammer that can see into the future" is a breach (precognition) of psychology. However, Boyer and Ramble have drawn the further distinction between breaches and "breach + transfers," which they argue is the application of a counterintuitive property (breach) to an inappropriate ontological category (transfer), or more simply, a transfer of a breach. Thus, for them, a "hammer that can see into the future" is a breach (precognition) plus a transfer

Once again, we asked 23 hypothesis-blind adults to provide independent ratings of the violation types by indicating whether they believed each new counterintuitive property described in Experiment 2 represented a violation of psychology, biology, or physics. As in Experiment 1, adult ratings of violation type (psychology, biology, or physics) were frequently inconsistent, and a majority consensus was achieved for only three of the six new counterintuitive properties introduced in Experiment 2. Notably, disagreement was once again most pronounced for the two properties we designated as violations of psychology, and also for one violation of physics. Paralleling the findings from Experiment 1, adults tended to categorize our proposed violations of psychology nearly equally as violations of psychology or biology. We suggest that categorization of these types of violations may be ambiguous because they simultaneously entail assumed violations of biology (i.e., it is impossible to experience psychological states without also

---

(psychological property attributed to an inanimate artifact typically lacking a psychology). Experiment 2 included two "breach + transfers" ("icicle/rock that could give birth to a teapot" and "hammer/rake that could see into the future") as defined by Boyer and Ramble (2001). To determine if there were recall differences between these "breach + transfers" and the "pure" breaches in our studies that contained no embedded transfer, we compared the average number of times children recalled nouns associated with each type of descriptive clause. Across children, nouns associated with "breach + transfer" descriptions and nouns associated with "pure" breach descriptions were recalled nearly equally (14 times compared to 14.7 times). Consequently, for the present studies, we use the term "breach" to refer to any counterintuitive feature that violates the assumed intuitive properties of members of any ontological category, without drawing a distinction between breaches that contain an embedded transfer and those that do not.

possessing the necessary underlying biological machinery, a brain, that renders this capacity possible).

As in Experiment 1, children were assigned to hear one of two versions of the fictional narrative and we counterbalanced whether nouns were paired with an intuitive or counterintuitive description across the two story versions. The target items and descriptions used in both story versions are presented in Table 3. The design and procedure for both the immediate and delayed recall tasks in Experiment 2 were in all other ways identical to those of Experiment 1. The same experimenter from Experiment 1 ran all participants in Experiment 2.

----------Insert Table 3 about here ----------

3.3. Results

3.3.1. Recall coding

The primary experimenter and a hypothesis-blind research assistant independently recorded and coded a transcript of each child's responses on the immediate recall task. The research assistant generated transcripts from video recordings for 95% of children's lab sessions. For the one session (5%) that was missing a video recording, the research assistant used the primary experimenter's live-recorded transcript to code children's responses. The research assistant coded children's delayed recall using the primary experimenters' live-recorded transcripts of the delayed recall task phone calls. Coders used the criteria for accurate recall established in Experiment 1 to code noun recall and overall target item recall using the 3-point coding system. Because inter-coder reliability was very high (98.3% agreement on all recall points awarded), the primary

experimenter's judgments were used in the minority of cases in which there was disagreement.

3.3.2. Immediate recall task

On average, children recalled 58.8% of the target item nouns (7.05 of 12 nouns) during the immediate story recall task. They scored an average of 15.6 recall points out of a possible total of 36 points, with an average 8.55 points corresponding to recall of descriptive clauses. A repeated measures ANOVA revealed a significant main effect of target item type on noun recall, $F(1, 19) = 18.97$, $p < .001$, $\eta^2 = .50$. Children recalled nouns with counterintuitive properties significantly more than nouns whose properties were entirely intuitive. On average, children accurately recalled 72.5% of the minimally counterintuitive nouns ($M = 4.35$ out of 6, $SD = 1.39$), but only 45.0% of intuitive nouns ($M = 2.7$ out of 6, $SD = 1.30$). There was a significant effect of item type on overall target item point recall (nouns + descriptive clauses), $F(1, 19) = 13.90$, $p = .001$, $\eta^2 = .423$. On average, children scored 9.65 out of a total of 18 possible recall points for counterintuitive items ($SD = 3.80$), but only 5.95 points for intuitive items ($SD = 3.39$). Thus, counterintuitive target items containing two violations of intuitive ontological expectations were recalled more frequently and in greater detail than intuitive target items. This effect was larger (though not statistically significantly so) for Experiment 2 than for Experiment 1. Across all subjects in Experiments 1 and 2, target item nouns were recalled more often when paired with counterintuitive descriptions (72.5%, 87 out of 120 times) relative to when those same nouns were paired with intuitive descriptions (45.0%, 54 out of 120 times). In addition, four of the top five most frequently recalled target item nouns were paired with counterintuitive descriptions.

There were no significant differences in the rate at which children recalled either zero or one adjectival clause in conjunction with successfully recalled intuitive and counterintuitive

target nouns. However, as in Experiment 1, children on average recalled two adjectives in conjunction with successfully recalled nouns approximately twice as often for counterintuitive nouns than for intuitive nouns (1.75 times compared to .95 times), $F(1, 19) = 6.91$, $p = .017$, $\eta^2 = .27$. Thus, children were significantly better at recalling all of the associated detail provided in the narrative for counterintuitive concepts than for intuitive concepts.

3.3.3. Delayed recall task

Of the 20 children who participated in the first lab session, 15 were successfully contacted one week later to complete the delayed recall task. Within subjects, delayed recall was highly correlated with immediate recall (nouns: $r = .78$, $p = .001$; points: $r = .85$, $p < .001$). However, children who participated in both the immediate and delayed recall tasks recalled significantly more target items and recalled them in greater detail during the immediate recall task compared to the delayed recall task (nouns: $F(1, 14) = 12.62$, $p = .003$, $\eta^2 = .47$; points: $F(1,14) = 21.60$, $p < .001$, $\eta^2 = .61$). On average, children who participated in both the immediate and delayed recall tasks recalled 6.93 of the possible 12 target item nouns during the initial lab visit, with recall dropping to 5.87 nouns one week later. These children scored an average of 15.0 out of 36 possible recall points during the immediate recall task, but only 11.20 points after one week.

One week after hearing the story, children recalled an average of 48.9% of the target nouns (5.87 of 12). There was a significant main effect of target item type on noun recall, $F(1,14) = 33.98$, $p < .001$, $\eta^2 = .71$. Children accurately recalled 68.9% of the counterintuitive nouns ($M = 4.13$ out of 6, $SD = 1.46$) but only 28.8% of the intuitive nouns ($M = 1.73$ out of 6, $SD = .59$). There was also a significant effect of item type on overall target item point recall, $F(1, 14) = 23.69$, $p < .001$, $\eta^2 = .63$. Children scored an average of 7.80 out of a total of 18 possible recall

points for counterintuitive items (*SD* = 3.26), but only 3.4 points for intuitive items (*SD* = 1.55). Fig. 4 displays differences in overall counterintuitive versus intuitive target item recall points (nouns + descriptive clauses) for both recall tasks. As in the immediate recall task, across all children, target nouns were recalled more often when paired with counterintuitive descriptions (68.9%, 62 out of 90 times) relative to when those same nouns were paired with intuitive descriptions (28.9%, 26 out of 90 times). In addition, across all children, the top five most frequently recalled target item nouns were all paired with counterintuitive descriptions.

----------Insert Figure 4 about here ----------

We performed paired samples t-tests on difference scores between children's recall for counterintuitive versus intuitive target items to detect differences in the magnitude of these effects between the immediate and delayed recall tasks. There were no differences for either nouns $t(14) = -1.57$, $p = .14$, nor recall points, $t(14) = 0.0$, $p = 1.0$. Thus, there was no observed differential decay of memory for counterintuitive concepts containing two violations of intuitive ontology and intuitive concepts over the course of one week.

3.4. Experiment 2 discussion

Children who listened to a fictional narrative containing both intuitive concepts and counterintuitive concepts containing two violations of intuitive ontological expectations recalled the counterintuitive concepts more frequently and in greater detail both during an immediate recall task and also one week later. The results of Experiment 2 offer evidence that for children,

concepts containing two distinct, non-overlapping violations of intuitive ontology are highly memorable. Accordingly, the next experiment compared children's memory for intuitive concepts and for concepts that violate intuition even more strongly.

4. Experiment 3: Children's memory for intuitive concepts vs. conceptions with three counterintuitive violations.

4.1. Participants

Participants were 8 boys and 13 girls, ranging in age from 7 to 9 years of age (M = 8.34 years, SD = .96, range = 7.01-9.96). All children were English-speakers, though 33% of children were raised in multilingual homes. We collected information about religious affiliation and religiosity for 61.9% of children (N = 13). These children were 69.2% Christian, 15.4% Jewish, 7.7% Muslim, and 7.7% Other. Most children came from fairly religious households in which parents felt it was quite important to raise their children in a religious tradition (modal score of 6 on the 1-7 scale). Most children (N = 9) attended religious services either every week or more often. In all other ways, the participants in Experiment 3 were equivalent to those who participated in the previous two experiments. Four parents requested being present in the experimental room to observe the immediate recall task.

4.2. Materials

The fictional story used in Experiment 3 was identical to the story used in the previous two

experiments except that the target items were lengthened in order to test the effect of three violations of intuitive ontological expectations on concept recall. We added a third adjectival clause to both the counterintuitive and intuitive items used in Experiment 2. New adjectival clauses were matched in approximate length (number of words) for corresponding counterintuitive and intuitive descriptions for all target nouns, and when grammatically necessary differed by no more than two words. All counterintuitive target items consisted of two breaches and one transfer.

For intuitive concepts, we included a third intuitive adjectival clause describing a property that adhered to each target noun's default domain-level properties (i.e., entailed no violation of intuitive ontology). We used the intuitive clauses originally paired with counterintuitive items in Experiment 1 to construct the third intuitive clause for intuitive items in Experiment 3, and made minor grammatical and descriptive alterations in order to achieve symmetrical word lengths.

For counterintuitive items, the third counterintuitive adjectival clause described a violation of the last remaining type (psychology, biology, or physics) not previously included in the counterintuitive items from Experiment 2. Thus, all counterintuitive target items contained a single violation in domains of psychology, biology, as well as physics. As in Experiment 2, we chose to use non-overlapping violations to ensure that children would process them distinctly, and not as complementary elaborations on the same general type of violation. In addition, this design allowed for a better test of whether counterintuitive items containing three non-overlapping violations of intuitive ontology prevent relevant inferences, and so exceed a cognitive threshold for memorability.

We asked 23 hypothesis-blind adults to provide independent ratings of violation type (psychology, biology, physics) for each of the new counterintuitive adjective clauses added to the

counterintuitive target items. In contrast to Experiments 1 and 2, adult ratings of violation type (psychology, biology, or physics) were consistent, and a majority consensus was achieved for all six new counterintuitive properties. Notably, both violations of psychology in Experiment 3 attributed psychological properties to beings that very clearly possess a biology (a lizard/rat and a flower/coconut tree). This may have mitigated the ambiguity in Experiments 1 and 2 in which psychological properties were attributed to inanimate objects (i.e., mailbox/stop sign, icicle/rock, and mango/banana), thereby potentially inviting confusion regarding the overlap between psychological and biological traits.

Once again, children were assigned to hear one of two versions of the fictional narrative, and we counterbalanced whether nouns were paired with either an intuitive or counterintuitive description across the two story versions. The target items and descriptions used in both story versions are presented in Table 4. The design and procedure for both the immediate and delayed recall tasks in Experiment 3 were in all other ways identical to those of the previous two experiments. The same experimenter from Experiments 1 and 2 ran all participants in Experiment 3.

----------Insert Table 4 about here ----------

4.3. Results

4.3.1. Recall coding

The primary experimenter and a hypothesis-blind research assistant independently recorded a transcript of each child's responses on the immediate recall task for 100% of the participant lab

sessions. Target item recall was coded both for noun recall alone and also using the 3-point coding system employed in the previous two experiments, using the same criteria for accurate recall. The research assistant coded children's delayed recall using the primary experimenter's live-recorded transcripts of the delayed recall task phone calls. Because inter-coder reliability was very high (98.4% agreement on all recall points awarded), the primary experimenter's judgments were used in the minority of cases in which there was disagreement.

4.3.2. Immediate recall task

On average, children recalled 56.3% of the target item nouns (6.76 of 12 nouns) during the immediate story recall task. This was equivalent to the percentage of nouns recalled in Experiments 1 (56.8%) and 2 (58.8%), despite the fact that nouns in Experiment 3 were paired with longer and more detailed descriptions. Children scored an average of 14.67 recall points out of a possible total of 36 points, with an average 7.90 points corresponding to recall of descriptive clauses. Thus, the overall amount of detail recalled about the target items as assessed by the point coding system was equivalent to Experiments 1 (14.1 recall points) and 2 (15.6 recall points). Therefore, the fact that the narrative used in Experiment 3 was slightly longer than the narratives used in Experiments 1 and 2 did not influence the overall amount of information children recalled. In Experiment 3, children rarely recalled all three adjectival clauses in association with either the intuitive (M = .33 out of 6 or 5.5%, SD = .58) or counterintuitive (M = .24 out of 6 or 4.0%, SD = .54) nouns.

A repeated measures ANOVA revealed no effect of target item type on noun recall, $F(1, 20) = .88$, $p = .36$. Nouns with three counterintuitive properties were not recalled significantly better than nouns with three intuitive properties. On average, children accurately recalled 59.5%

of the counterintuitive nouns (M = 3.57 out of 6, SD = 1.66) and 53.2% of intuitive nouns (M = 3.19 out of 6, SD = 1.21). Similarly, there was no effect of item type on overall target item point recall (nouns + descriptive clauses), $F(1, 20) = .036$, $p = .85$. On average, children scored 7.43 out of a total of 18 possible recall points for counterintuitive items (SD = 4.02) and a comparable 7.24 points for intuitive items (SD = 2.97). Across subjects, target nouns were not recalled differently when paired with either counterintuitive or intuitive descriptions. Unlike in Experiments 1 and 2, recall for counterintuitive target items, which contained three violations of intuitive ontological expectations in Experiment 3, was indistinguishable from recall for intuitive target items in terms of both frequency and detail.

3.3.3. Delayed recall task

Of the 21 children who participated in the initial lab session, 17 were successfully contacted one week later to complete the delayed recall task. Within participants, delayed recall was highly correlated with immediate recall (nouns: $r = .86$, $p < .001$; points: $r = .88$, $p < .001$). Children who participated in both the immediate and delayed recall tasks recalled target nouns only marginally better during the immediate recall task compared to the delayed recall task, $F(1,16) = 3.77$, $p = .07$, but they recalled those nouns in significantly greater detail during the immediate recall task than in delayed recall task, $F(1, 16) = 20.75$, $p < .001$, $\eta^2 = .57$). On average, children who participated in both recall tasks recalled 6.88 of the possible 12 target item nouns during the initial lab visit and 6.35 nouns one week later. These children scored an average of 14.82 out of 36 possible recall points during the immediate recall task, but only 11.82 points after one week.

One week after hearing the story, children recalled an average of 52.9% of the target nouns

(6.35 of 12). As in the immediate recall task, there was no within-subjects effect of target item type on noun recall, $F(1, 16) = .74$, $p = .40$. Children accurately recalled 55.8% of the counterintuitive nouns ($M = 3.35$ out of 6, $SD = 1.41$) and 50.0% of the intuitive nouns ($M = 3.0$ out of 6, $SD = 1.37$). There was also no effect of item type on overall target item point recall, $F(1, 16) = .046$, $p = .83$. Children scored an average of 6.0 out of a total of 18 possible recall points for counterintuitive target items ($SD = 3.59$), and 5.82 for intuitive target items ($SD = 2.92$). Across subjects, target nouns were not recalled differently when paired with either counterintuitive or intuitive descriptions. In sum, there were no detectable differences in recall for counterintuitive and intuitive concepts one week after children heard the story. Fig. 5 displays counterintuitive and intuitive target item point recall (nouns + adjectival clauses) for both the immediate and delayed recall tasks.

----------Insert Figure 5 about here ----------

3.4. Experiment 3 discussion

Children who listened to a fictional narrative containing both intuitive concepts and concepts presenting three violations of intuitive ontological expectations recalled both types of concepts equally well during an immediate recall task and also one week later. In contrast to Experiments 1 and 2, nouns that were paired with counterintuitive descriptions containing three violations were recalled neither more frequently nor in greater detail than intuitive concepts. We conclude that concepts that contain three distinct, non-overlapping violations of psychology, biology, and physics do not demonstrate the recall advantage characteristic of concepts containing

a limited number of violations (one or two) because they cross the threshold of cognitive optimality for children. Following Boyer's model (1994, 2000), the counterintuitive concepts from Experiment 3 therefore appear to be too counterintuitive for children. These results support the hypothesis that conceptual counterintuitiveness is curvilinearly related to concept recall. Concepts containing a minimal or moderate number of violations of intuitive ontology are recalled better than concepts containing no violations, but concepts with too many violations are not.

4. Comparison of results across Experiments 1, 2, and 3

Because Experiments 1 and 2 indicated that concepts containing either one or two violations of intuitive ontological expectations are both minimally counterintuitive, we contrasted recall for minimally counterintuitive and maximally counterintuitive concepts by comparing collapsed data from Experiments 1 and 2 with data from Experiment 3. A 2x2 repeated measures ANOVA with concept type (counterintuitive versus intuitive) as the within-subjects factor and experiment (1 + 2 versus 3) as the between-subjects factor yielded a significant concept type by experiment interaction effect for recall points, $F(1, 81) = 7.04$, $p = .01$, $\eta^2 = .08$ (Fig. 6), but not for nouns, $F(1, 81) = 2.63$, $p = .11$, on the immediate recall task. We performed an identical 2x2 repeated measures ANOVA for the delayed recall tasks, in order to compare recall of minimally counterintuitive and maximally counterintuitive concepts one week after children heard the story. Once again, there was a significant concept type by experiment interaction effect for recall points, $F(1, 51) = 8.29$, $p = .005$, $\eta^2 = .11$ (Fig. 7), but not for nouns, $F(1, 51) = 2.4$, $p = .13$.

----------Insert Figure 6 about here ----------

----------Insert Figure 7 about here ----------

5. General discussion

We have shown that 7-9-year old children who listened to a story including both intuitive and minimally counterintuitive concepts containing one (Experiment 1) or two (Experiment 2) violations of intuitive ontological expectations recalled the minimally counterintuitive concepts significantly more and in greater detail than the intuitive concepts, both immediately after hearing the story and one week later. Remarkably, this recall advantage for minimally counterintuitive concepts was so robust that it rendered nouns paired with minimally counterintuitive descriptions more memorable than nouns paired with intuitive descriptions, despite the fact that children frequently failed to recall the actual descriptors. Because the nouns that were used, by themselves, are neither counterintuitive nor intuitive, this finding suggests that nouns embedded in an minimally counterintuitive conceptual formulation possess a cognitive advantage at the level of encoding. However, counterintuitive concepts containing three violations of expectation showed no recall advantage over intuitive concepts with the same number of descriptors (Experiment 3), suggesting that one or two violations of expectation may be a cognitive optimum for children, in the sense of being more inferentially rich and so more memorable, but that any increase in counterintuitiveness beyond this threshold diminishes memory for target concepts.

The results of the present set of studies therefore suggest that, in children, conceptual counterintuitiveness is curvilinearly related to concept recall. Like adults, children demonstrate superior immediate as well as delayed recall of minimally counterintuitive concepts, but not

extremely counterintuitive concepts, compared to intuitive concepts embedded in the context of a fictional narrative. Children therefore seem to be sensitive to conceptual counterintuitiveness, an important developmental continuity in specific cognitive biases that account for supernatural beliefs.

We explain these differential recall results not in terms of the inherent memorability in children of a specific number of violations *per se* (which would only push the question back as to why this specific number of items is special), but in terms of relevance theory (Sperber & Wilson, 1995). For a concept to be memorable, it has to be relevant; to be relevant, it has to combine with prior knowledge in some domain to generate a rich set of inferences. Crucially, relevant inferences can be made by children in natural domains of reasoning (physics, biology, psychology) that are *not* violated. In Experiment 1, there were two possible domains that were available for rich sets of inferences about the default properties and behaviors of each entity labeled by a noun, and in Experiment 2, there was one available natural domain for inferences about each noun. The availability of relevant inferences across either one or two non-violated natural domains allowed for superior recall in these two experiments, compared to their intuitive counterparts. Although Experiments 1 and 2 varied the number of counterintuitive violations described, both yielded similar results because, in both Experiments 1 and 2, counterintuitive target items were more salient than entirely intuitive items, but in neither did they cross the threshold of cognitive optimality beyond which a loss in inferential productivity impoverishes target item recall. However, this was not the case for Experiment 3, wherein the maximally counterintuitive conditions had no natural domains in which properties were not violated, and in which inferences could be produced. This approach explains why the relative recall advantage over intuitive concepts that exists for minimally counterintuitive concepts disappears when they

become so counterintuitive that they license no productive inferences. Although attention is heightened in theory by the addition of a third violation, that is more than offset by the disappearance of any domain for possible relevant inferences. This approach may also explain why children recalled intuitive items somewhat *better* in Experiment 3 than in Experiments 1 and 2. We suggest that this is because three descriptors are better than two or one when the descriptors are intuitive, and relevant inferences can be produced.

We note that although target items in Experiment 3 were longer than those used in Experiments 1 and 2, differences in target item length are unlikely to account for the pattern of recall results obtained across all three experiments. First, children in Experiment 3 recalled the same overall amount of story information as did children in Experiments 1 and 2, despite the longer length of target items. Children in Experiment 3 almost never recalled all three adjectival clauses associated with either intuitive or counterintuitive nouns (only 4.8% of the time), but instead, nearly always dropped one of the target descriptors. If the longer length or additional detail contained in these target items caused children to fail to fully encode, on average, one of the descriptors per target item, then children would have effectively transformed counterintuitive items with three descriptors into items with two counterintuitive features, and intuitive items with three descriptors into items with two intuitive features. If this were true, on the basis of results from Experiment 2, we would have expected to see a robust recall advantage for counterintuitive items with two counterintuitive descriptors over intuitive items with two intuitive descriptors. As this was not the case, it is unlikely that differences in target length can account for the pattern of results obtained.

We also point out that intuitive target items in Experiment 3 were recalled *better* than intuitive target items in Experiments 1 and 2, but this same pattern was not true of

counterintuitive target items. This finding suggests that differences in recall across the three experiments cannot straightforwardly be explained by differences in target item length, as increased length had a different effect on intuitive item recall than it did on counterintuitive item recall. Instead, we believe that the most parsimonious account of the obtained results is that three descriptors are better than one or two when the descriptors are intuitive, and relevant inferences can be produced. However, this is not true for counterintuitive items containing three distinct violations in the domains of psychology, biology, and physics because these target items possess no natural domains in which properties are not violated and in which inferences could be produced. The disappearance of any domain for possible relevant inferences eliminated the recall advantage of counterintuitiveness for these items. We therefore interpret these results to suggest that recall differences across the three experiments were due to differences in encoding that were a function of the conceptual content of target items and not merely of target item length. However, future empirical work would benefit from more systematically controlling for target item length while varying conceptual content.

Given that religious thought, feeling and behavior are by-products of evolved brain functions (Boyer, 2001; Atran, 2002; Boyer & Liénard, 2006; Bloom, 2007), it is reasonable that children would be equipped with some of the same cognitive predispositions for religious beliefs and practices that adults possess. The finding that certain aspects of the relevant underlying evolved cognitive mechanisms that give rise to religious beliefs and practices are active from an early age may help to explain not only why religious thoughts and behaviors are human universals, but also why there are so many cross-culturally recurrent features of religions documented by anthropologists and historians of religion (Tylor, 1871; Brown, 1991; Eliade, 1959; Boyer, 2001; Atran, 2002). Consistent with epidemiological models of cultural

representations (e.g., Sperber, 1985, 1996; Boyer, 1994), the early minimally counterintuitive cognitive bias we report would also help explain the cultural transmission of both religious and non-religious literatures that feature supernatural elements, as in poetry, mythological stories and folktales. Interestingly, the recall advantage for minimally counterintuitive ideas appears not only to operate at the level of individual concepts, but also at the level of narratives as a whole (Norenzayan et al., 2006). For example, Norenzayan et al. (2006) found that Grimm Brothers' folktales that contained two violations embedded within their plots, but neither more nor fewer, were most culturally successful. This parallels the results of the present experiments, which show that, for children, two counterintuitive elements seems to be a cognitive threshold beyond which the recall advantage for counterintuitive ideas over entirely intuitive ideas diminishes.

Importantly, our findings reveal that children's preferential recall of minimally counterintuitive concepts is unaffected by their exposure to counterintuitive religious concepts in the home environment, and is therefore unlikely to be a product of learning. Specifically, two measures of familial religiosity that are reasonable indicators of exposure to counterintuitive religious ideas—parents' desire to raise their child in a religious tradition and frequency of attendance at religious services—were uncorrelated with children's performance on the recall tasks[5]. This result suggests that preferential recall of minimally counterintuitive concepts is a reliably developing representational bias that is relatively insensitive to environmental differences

[5] We note that religious concepts are only one potential source of counterintuitive ideas. The ubiquity of fantasy themes in children's literature, television, and games, suggests that these other sources may play a proportionally much greater role than religion in exposing children to counterintuitive ideas.

in the frequency of exposure to concepts that fit this particular conceptual template.

One perplexing finding concerns the sometimes inconsistent nature of independent adult ratings of our assigned minimally counterintuitive property violation types (biology, psychology, physics), especially for our designations of violations of psychology, which constituted the main area of disagreement. It seems that when disagreement occurred, it was over whether inanimate objects with mental states should be classified as a biological or a psychological violation, and participants tended to choose biology. This may have been because participants thought that the capacity to have mental states requires underlying biological machinery, or could be a result of the commonality in vitalist and animistic/anthropomorphic intuitions, both of which derive from psychological essentialism and intuitive dualism (Gelman, 2003; Bloom, 2004).

This study also suggests the importance of some open questions that should be investigated further. Future work should characterize in more detail the immediate as well as delayed recall advantage, observed in Experiment 1, that children had for violations that were breaches, rather than transfers, of intuitive expectations. Specifically, is this advantage due to the psychological effects of breaches *per se* (perhaps due to superior encoding, less blocking of inference generation)? Potentially, breaches are more surprising and salient than transfers because they refer to unfamiliar properties that are untrue of members of any ontological category rather than properties that are familiar but only inappropriately applied to members of the wrong ontological category. This heightened saliency may render breaches more attention-demanding than transfers, and therefore better encoded and recalled. Alternatively, the observed recall advantage for breaches in the present experiments could be due to the fact that violations of intuitive physics happened to be more numerous among breaches. A more valid future experiment should test recall after distributing breaches and transfers equally across the three violation types (physics,

biology, psychology). A further open question is whether violations are more memorable when they transgress a greater ontological distance between domains. In addition, as predicted by relevance theory, might multiple violations of different domains be more counterintuitive than the same number of violations within one domain? Furthermore, might factors of ontological distance and violation type (breach/transfer) have some kind of interaction? Specifically, among either breaches or transfers alone, are there unique recall advantage hierarchies for particular ontological domain violations?

The present study has focused on memory and measuring item recall in the investigation of minimally counterintuitive concepts in children. A number of investigations may provide further support for this bias in the cognition of children. First, the minimally counterintuitive effect is hypothesized to be the outcome of more than just differential memory. Specifically, attention, as well as memorability, is predicted to interact in the production of minimally counterintuitive inferences (Boyer, 1994, 2000, 2001). Further work in children (and adults) should therefore move beyond measuring recall alone, and include tasks to independently measure attention and inference-making in response to intuitive, minimally counterintuitive, and maximally counterintuitive concepts.

Second, in adults, there is evidence that minimally counterintuitive concepts are especially attractive choices for subjects when they are asked to create novel stories, compared to intuitive and maximally counterintuitive concepts (Tweney, Upal, Gonce, Slone, & Edwards, 2006). The differential presence of minimally counterintuitive concepts in the spontaneous *production* of stories in adults complements evidence on the preferential recall of minimally counterintuitive concepts. Whether children also prefer to spontaneously create stories with minimally counterintuitive concepts, and whether in iterations of transmission minimally counterintuitive

concepts are differentially preserved could be investigated. These studies would complement the evidence presented in this investigation, and would also help explain the particular magico-religious nature of literature consumed by children (Grimm & Grimm, 1869/2003).

Third, in delayed recall tasks, consistent with our hypothesis, there was a significant effect of item type (counterintuitive vs. intuitive) on overall target item recall. However, in contrast to the findings of some studies of adults (Gonce et al., 2006), we observed no overall differential decay of minimally counterintuitive and intuitive concept memory over the course of one week. Minimally counterintuitive concepts were recalled better than intuitive concepts on both the immediate and delayed recall task, but the magnitude of this effect did not vary as a function of time. It is possible that relatively high overall recall rates during the delayed task (in relation to immediate recall rates) partially masked the differential fitness of these two conceptual categories over time, because the one-week delay was not long enough to reveal the differential decay of the different types of concepts (adult studies have used a delay of 3 months, Barrett & Nyhof, 2001). Although a delay of one week was chosen in the current study to avoid potential recall floor effects in which children's memory for the story was extremely low, further investigations could use longer delays to map the full scale of recall decay.

In sum, the results of these three experiments provide evidence that the cognitive bias for minimally counterintuitive ideas accounting for the cultural success of supernatural conceptual templates in adults is present and active early in human development, near the start of formal religious instruction. This finding supports a growing literature suggesting that diverse, early-emerging, evolved psychological biases predispose humans toward religious beliefs and practices whose primary form as well as content is not arbitrarily derived from custom or the social environment.

Acknowledgements

We thank Ellyn Schmidt for her assistance with data collection and Christine Size, Stephanie O'Connell, and Shelley Pearson for their help with coding data. This research was supported by a grant from NIH (5R01 HD23103-27) to E.S.S.

References

Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. New York: Oxford University Press.

Atran, S., & Norenzayan, A. (2004). Religion's evolutionary landscape: Counterintuition, commitment, compassion, communion. *Behavioral and Brain Sciences, 27* (6)*,* 713-770.

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

Barrett, J. L. (2004). *Why would anyone believe in God?* Walnut Creek, CA: AltaMira Press.

Barrett, J.L., & Nyhof, M.A. (2001). Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *Journal of Cognition and Culture*, *1* (1)*,* 69-100.

Beardsworth, E., & Bishop, D. (1994). Assessment of long-term verbal memory in children. *Memory, 2* (2), 129-148.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.

Bloom, P. (2007). Religion is natural. *Developmental Science*, *10* (1), 147-151.

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

Boyer, P. (2000). Functional origins of religious concepts: Conceptual and strategic selection in evolved minds*. Journal of the Royal Anthropological Institute, 6* (2), 195-214.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences, 7* (3), 119-124.

Boyer, P., & Ramble, C. (2001). Cognitive templates of religious concepts: Cross-cultural evidence for recall of counter-intuitive representations. *Cognitive Science, 25* (4), 535-564.

Boyer, P., & Barrett, H. C. (2005). Evolved intuitive ontology: Integrating neural, behavioral and developmental aspects of domain-specificity. In David Buss (Ed.), *Handbook of evolutionary psychology* (pp. 96-118). New York: John Wiley.

Boyer, P., & Liénard, P. (2006). Why ritualized behavior? Precaution systems and action parsing in developmental, pathological, and cultural rituals. *Behavioral and Brain Sciences, 29* (6), 1-56.

Brown, D. E. (1991). *Human universals*. New York: McGraw-Hill.

Courage, M., & Cowan, N. (2008). *The development of memory in infancy and childhood*. (2nd ed.). New York: Psychology Press.

Eliade, M. (1959). *The sacred and the profane: The nature of religion*. New York: Harcourt Brace Jovanovich.

Gambrell, L.B., & Jawitz, P.B. (1993). Mental imagery, text illustrations, and children's story comprehension and recall. *Reading Research Quarterly, 28* (3), 264-276.

Gelman, S. (2003). *The essential child: Origins of essentialism in everyday life*. New York: Oxford University Press.

Goody, J. (1987). *The interface between the written and the oral*. Cambridge, UK: Cambridge University Press.

Gonce, L.O., Upal, M.A., Slone, D.J., & Tweney, R.D. (2006). Role of context in the recall of counterintuitive concepts. *Journal of Cognition and Culture, 6*, 521-547.

Graham, W. (1993). *Beyond the written word*: *Oral aspects of scripture in the history of religion*. New York: Cambridge University Press.

Grimm, J., & Grimm, W. (2003). *Grimm's fairy tales*. New York: Barnes & Noble Classics. (Originally published 1869).

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science, 30* (3), 531-553.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. New York: Viking Penguin.

Pyysiäinen, I., & Hauser, M. (2010). The origins of religion: Evolved adaptation or by-product? *Trends in Cognitive Sciences, 14* (3), 104-109.

Pyysiäinen, I., Lindeman, M., & Honkela, T. (2003). Counterintuitiveness as the hallmark of religiosity. *Religion*, *33* (4), 341-355.

Spelke, E., & Kinzler, K. (2007). Core knowledge. *Developmental Science*, *10* (1), 89–96.

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

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences, 8* (1)*,* 40-46.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.

Tweney, R.D., Upal, M.A., Gonce, L.O., Slone, D. J., & Edwards, K. (2006). The creative structuring of counterintuitive worlds. *Journal of Cognition and Culture, 6* (3-4), 483-498.

Tylor, E.B. (1871). *Primitive culture*. London: John Murray.

Upal, M.A., Gonce, L.O, Tweney, R.D., & Slone, D.J. (2007). Contextualizing counterintuitiveness: How context affects comprehension and memorability of counterintuitive concepts. *Cognitive Science, 31* (3), 415-439.

Wellman, H.M., & Gelman, S.A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43,* 337-375.

Zwaan, R.A. (1993). *Aspects of literary comprehension: A cognitive approach.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Appendix A.

Experiment 1, Story Version 1:

Once upon a time, a brother and a sister named Joe and Jessie moved with their parents to a new house on a new street that they had never seen before in a town just down the road from where they used to live. The kids were excited to explore their new home and to learn more about the neighborhood. So as soon as their boxes were unpacked, Joe and Jessie decided to go see what they could find. First, they climbed up a staircase and went into the attic, where they saw a **lizard** on the floor. This was a lizard that had a long, thin tail and could never die no matter how old it was. The kids left the attic and wandered to their parent's bedroom. In the bedroom, they saw a **hammer** lying on the carpet. The hammer had a wooden handle and could breathe well underwater. After leaving the bedroom, the kids continued on into the basement, where they noticed a **rake** on top of a table. The rake felt heavy to hold and was brown in color.

Growing bored of the house, the kids went outdoors into their new backyard. When they looked up, they saw an **icicle** just above their heads on the roof of the garage. The icicle felt cold to touch and liked to sing very loudly. The kids turned their heads to look at the street in front of their house and noticed a **stop sign** at the corner of the road. This was a stop sign that was covered with rust and was crying because it was sad. The kids skipped down the street and came across a garden that had a single **flower** in it. The flower moved from side to side in the wind, and it could disappear and reappear in a different spot in the garden. The kids finally reached the front yard of their closest neighbor's house. On the lawn, the kids spotted a **rat**. The rat ate insects off the ground and crawled around quickly on all four of its feet.

Joe and Jessie then decided to sneak behind their neighbor's house to see what they might find in the backyard. While they were there, they saw a **rock** on the shed next to the house. This was a rock that was thick and hard and looked shiny in the sunlight. The kids noticed that behind their neighbor's house was a small woody area. When they reached the woods, they saw a coco**nut tree** planted in the ground. The coconut tree had roots that went deep into the soil and needed lots of sunshine and water from the rain to grow. In the middle of the woods, the kids came across a **banana** lying on top of a pile of leaves. The banana was very fresh and ripe and turned invisible every few minutes.

The kids were tired of searching the woods, so they ran back out to the street. They were getting hungry now, because they had been exploring for a long time. So they decided to go home and make themselves a snack. On their way home, they saw a **mailbox** stuck in the ground. The mailbox was made of metal and had sharp edges along its corners. They were almost back at their house when they noticed a **mango** lying in the grass in front of their porch. This was a mango that had a bright yellow skin and smelled very fruity and delicious. Finally, the kids arrived at the front door of their house again. They went inside thinking that their new neighborhood was going to be a very interesting place to live.

Experiment 1, Story Version 2:

Once upon a time, a brother and a sister named Joe and Jessie moved with their parents to a new house on a new street that they had never seen before in a town just down the road from where they used to live. The kids were excited to explore their new home and to learn more about the neighborhood. So as soon as their boxes were unpacked, Joe and Jessie decided to go see what they could find. First, they climbed up a staircase and went into the attic, where they saw a **rat** on the floor. This was a rat that had a long, thin tail and could never die no matter how old it was. The kids left the attic and wandered to their parent's bedroom. In the bedroom, they saw a **rake** lying on the carpet. The rake had a wooden handle and could breathe well underwater. After leaving the bedroom, the kids continued on into the basement, where they noticed a **hammer** on top of a table. The hammer felt heavy to hold and was brown in color.

Growing bored of the house, the kids went outdoors into their new backyard. When they looked up, they saw a **rock** just above their heads on the roof of the garage. The rock felt cold to touch and liked to sing very loudly. The kids turned their heads to look at the street in front of their house and noticed a **mailbox** at the corner of the road. This was a mailbox that was covered with rust and was crying because it was sad. The kids skipped down the street and came across a garden that had a single **coconut tree** in it. The coconut tree moved from side to side in the wind, and it could disappear and reappear in a different spot in the garden. The kids finally reached the front yard of their closest neighbor's house. On the lawn, the kids spotted a **lizard**. The lizard ate insects off the ground and crawled around quickly on all four of its feet.

Joe and Jessie then decided to sneak behind their neighbor's house to see what they might find in the backyard. While they were there, they saw an **icicle** on the shed next to the house. This was an icicle that was thick and hard and looked shiny in the sunlight. The kids noticed that behind their neighbor's house was a small woody area. When they reached the woods, they saw a **flower** planted in the ground. The flower had roots that went deep into the soil and needed lots of sunshine and water from the rain to grow. In the middle of the woods, the kids came across a **mango** lying on top of a pile of leaves. The mango was very fresh and ripe and turned invisible every few minutes.

The kids were tired of searching the woods, so they ran back out to the street. They were getting hungry now, because they had been exploring for a long time. So they decided to go home and make themselves a snack. On their way home, they saw a **stop sign** stuck in the ground. The stop sign was made of metal and had sharp edges along its corners. They were almost back at their house when they noticed a **banana** lying in the grass in front of their porch. This was a banana that had a bright yellow skin and smelled very fruity and delicious. Finally, the kids arrived at the front door of their house again. They went inside thinking that their new neighborhood was going to be a very interesting place to live.

Appendix B.

Target Item Order Control Experiment Story:

Once upon a time, a brother and a sister named Joe and Jessie moved with their parents to a new house on a new street that they had never seen before in a town just down the road from where they used to live. The kids were excited to explore their new home and to learn more about the neighborhood. So as soon as their boxes were unpacked, Joe and Jessie decided to go see what they could find. First, they climbed up a staircase and went into the attic, where they saw a **lizard** on the floor. This was a lizard that had a long, thin tail and could never die no matter how old it was. The kids left the attic and wandered to their parent's bedroom. In the bedroom, they saw a **hammer** lying on the carpet. The hammer had a wooden handle and could breathe well underwater. After leaving the bedroom, the kids continued on into the basement, where they noticed a **rake** on top of a table. The rake felt heavy to hold and was brown in color.

Growing bored of the house, the kids went outdoors into their new backyard. When they looked up, they saw a **rock** just above their heads on the roof of the garage. This was a rock that was thick and hard and looked shiny in the sunlight. The kids turned their heads to look at the street in front of their house and noticed a **stop sign** at the corner of the road. This was a stop sign that was covered with rust and was crying because it was sad. The kids skipped down the street and came across a garden that had a single **coconut tree** in it. The coconut tree had roots that went deep into the soil and needed lots of sunshine and water from the rain to grow. The kids finally reached the front yard of their closest neighbor's house. On the grass, the kids spotted a **banana** lying on top of a pile of leaves. The banana was very fresh and ripe and turned invisible every few minutes.

Joe and Jessie then decided to sneak behind their neighbor's house to see what they might find in the backyard. While they were there, they saw an **icicle** on the shed next to the house. The icicle felt cold to touch and liked to sing very loudly. The kids noticed that behind their neighbor's house was a small woody area. When they reached the woods, the kids spotted a **rat**. The rat ate insects off the ground and crawled around quickly on all four of its feet. In the middle of the woods, the kids came across a mango lying on the ground. This was a **mango** that had a bright yellow skin and smelled very fruity and delicious.

The kids were tired of searching the woods, so they ran back out to the street. They were getting hungry now, because they had been exploring for a long time. So they decided to go home and make themselves a snack. On their way home, they saw a **flower** planted in the lawn. The flower moved from side to side in the wind, and it could disappear and reappear in a different spot in the lawn**.** They were almost back at their house when they noticed a **mailbox** stuck in the ground. The mailbox was made of metal and had sharp edges along its corners. Finally, the kids arrived at the front door of their house again. They went inside thinking that their new neighborhood was going to be a very interesting place to live.

Figure Captions

Fig. 1. Experiment 1 Recall of a Single Adjectival Clause for Counterintuitive Nouns.
** $p \leq .01$.

Fig. 2. Experiment 1 Breach vs. Transfer Recall.
* $p \leq .05$.
*** $p \leq .001$.

Fig. 3. Experiment 1 Point Recall.
*** $p \leq .001$.

Fig. 4. Experiment 2 Point Recall.
*** $p \leq .001$.

Fig. 5. Experiment 3 Point Recall.
** $p \leq .01$.

Fig. 6. Minimally Counterintuitive (Exp. 1 + Exp. 2) vs. Maximally Counterintuitive (Exp. 3) Immediate Point Recall.
** $p \leq .01$.

Fig. 7. Minimally Counterintuitive (Exp. 1 + Exp. 2) vs. Maximally Counterintuitive (Exp. 3) Delayed Point Recall.
** $p \leq .01$.

Table Captions

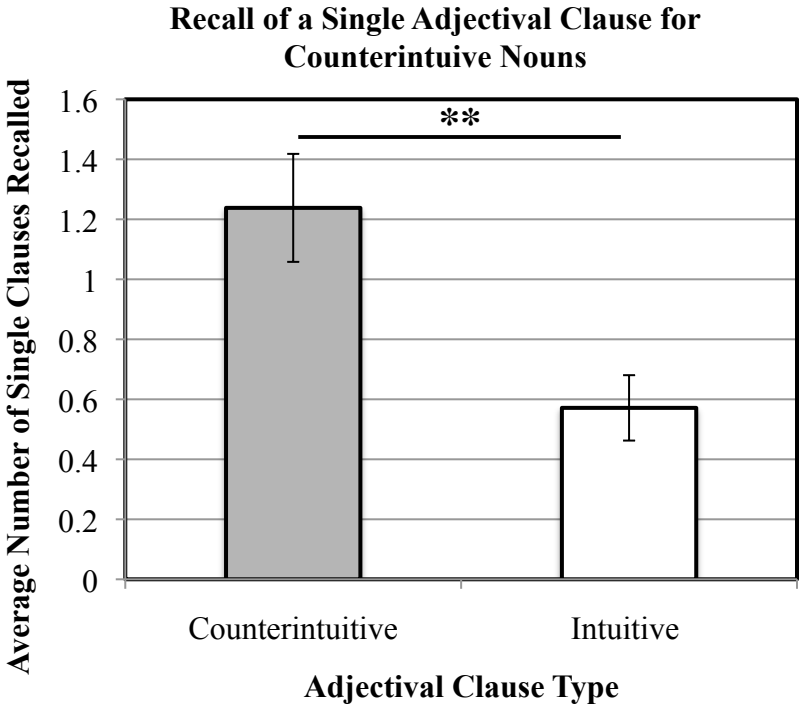

Fig. 1. Experiment 1 Recall of a Single Adjectival Clause for Counterintuitive Nouns.
** $p \leq .01$.

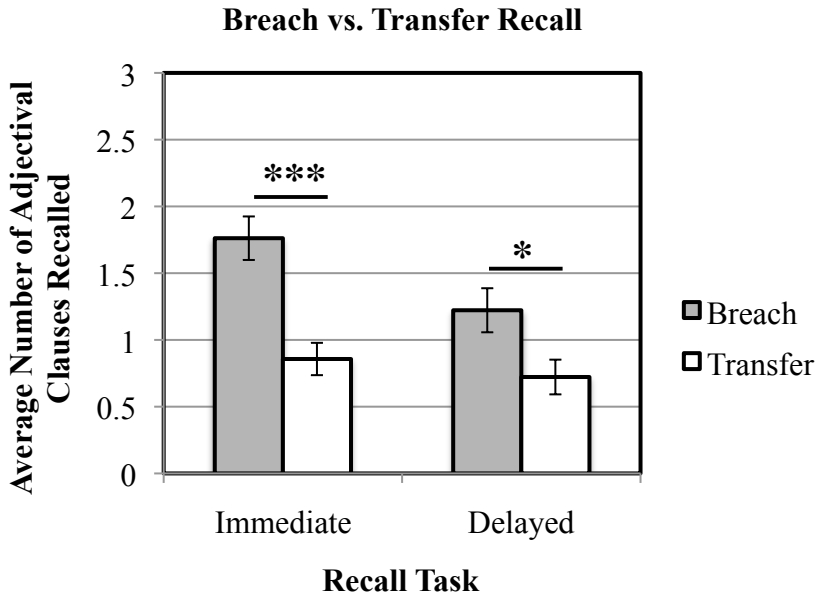

Fig. 2. Experiment 1 Breach vs. Transfer Recall.
* $p \leq .05$.
*** $p \leq .001$.

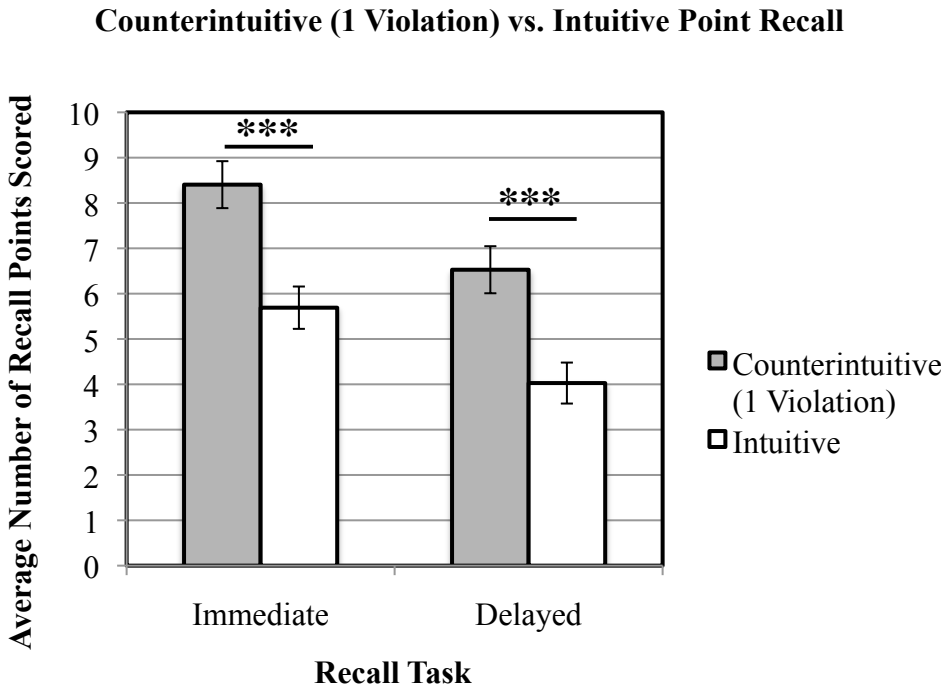

Fig. 3. Experiment 1 Point Recall.
*** $p \leq .001$.

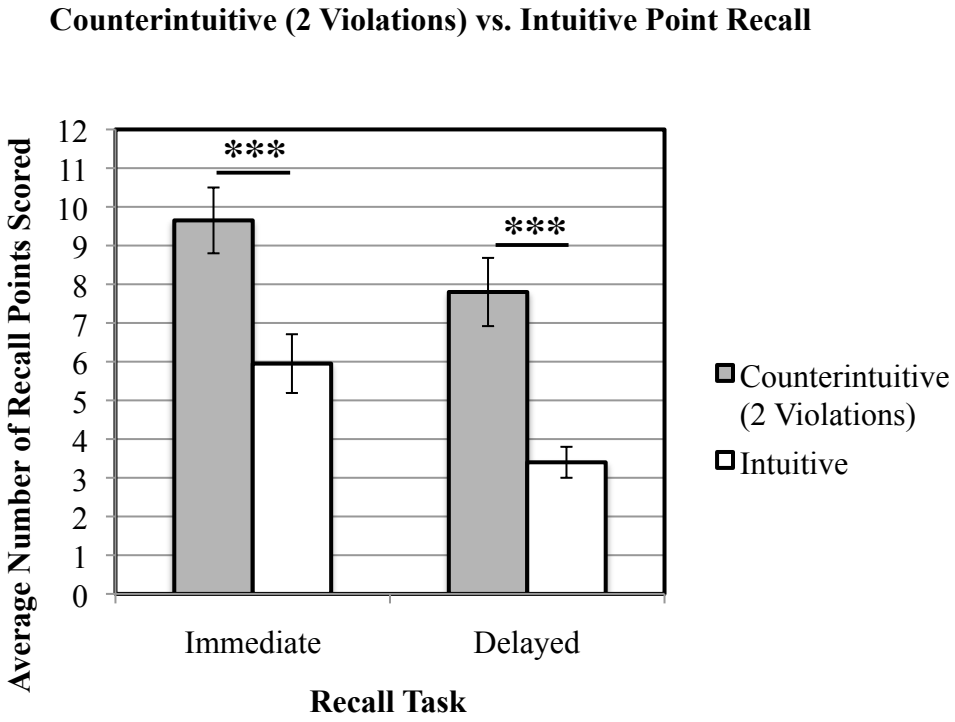

Fig. 4. Experiment 2 Point Recall.
*** $p \leq .001$.

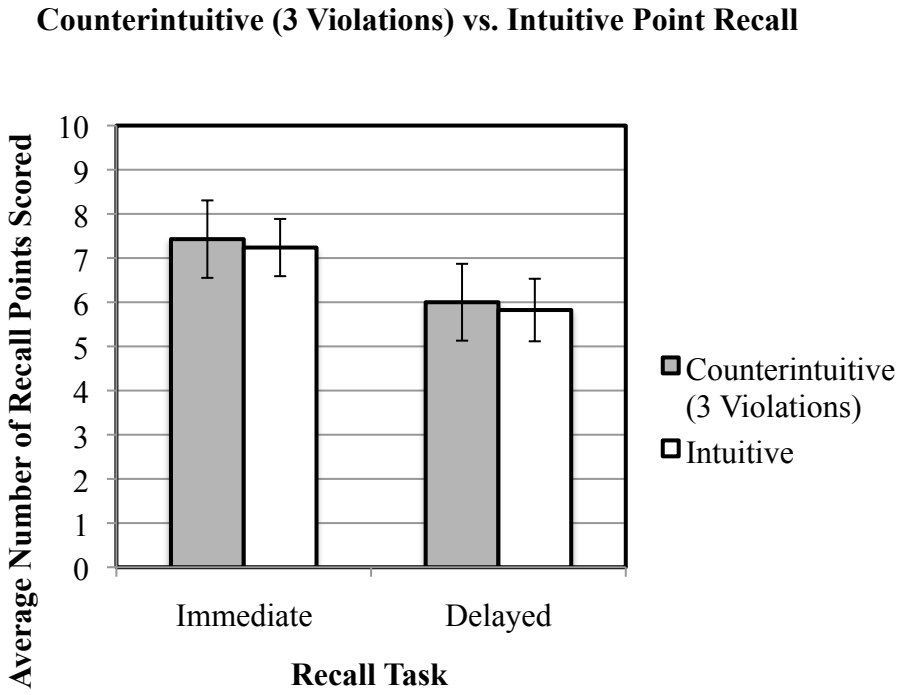

Fig. 5. Experiment 3 Point Recall.
** $p \leq .01$.

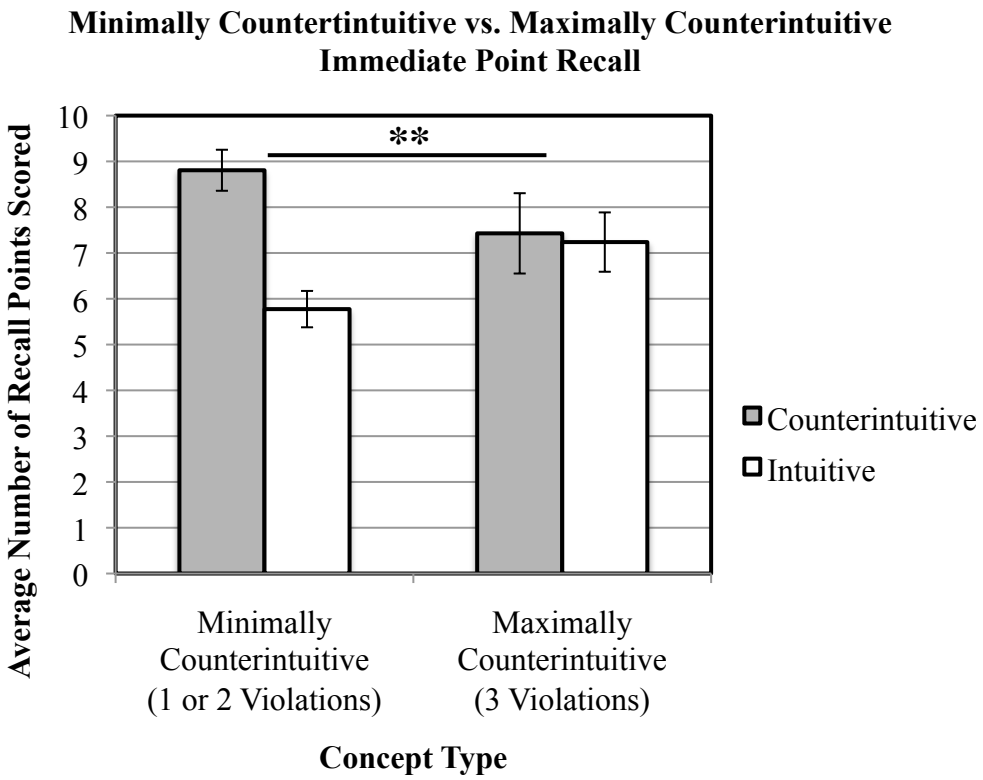

Fig. 6. Minimally Counterintuitive (Exp. 1 + Exp. 2) vs. Maximally Counterintuitive (Exp. 3) Immediate Point Recall.
** $p \leq .01$.

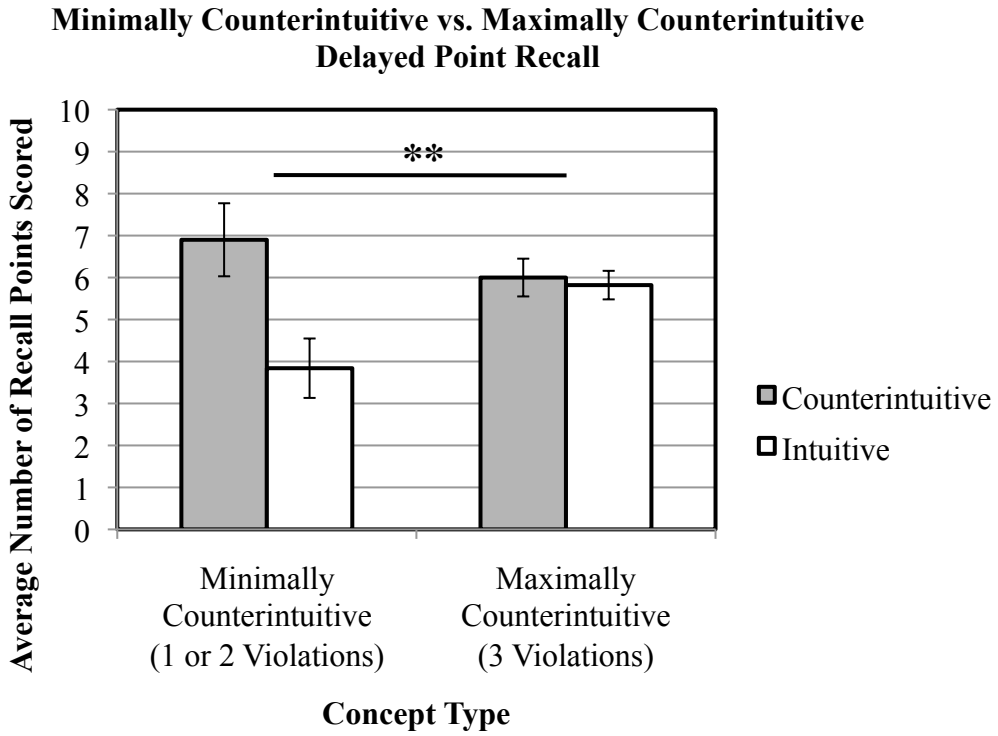

Fig. 7. Minimally Counterintuitive (Exp. 1 + Exp. 2) vs. Maximally Counterintuitive (Exp. 3) Delayed Point Recall.
** $p \leq .01$.

Table 1. Experiment 1 Target Items.

Target Noun Pairs	Ontological Category	Target Items		
		Intuitive		
1. lizard/rat	animal	that ate insects off the ground and crawled around quickly on all four of its feet		
2. flower/coconut tree	plant	that had roots that went deep into the soil and needed lots of sunshine and water from the rain to grow		
3. banana/mango	plant	that had a bright yellow skin and smelled very fruity and delicious		
4. icicle/rock	natural object	that was thick and hard and looked shiny in the sunlight		
5. hammer/rake	human-made artifact	that felt heavy to hold and was brown in color		
6. stop sign/mailbox	human-made artifact	that was made of metal and had sharp edges along its corners		
		Counterintuitive (1 violation)	**Breach/Transfer**	**Violation Type**
1. lizard/rat	animal	that had a long, thin tail and could never die no matter how old it was	Breach	Biology
2. flower/coconut tree	plant	that moved from side to side in the wind and could disappear and reappear in a different spot in the garden	Breach	Physics
3. banana/mango	plant	that was very fresh and ripe and turned invisible every few minutes	Breach	Physics
4. icicle/rock	natural artifact	that felt cold to touch and liked to sing very loudly	Transfer	Psychology
5. hammer/rake	human-made artifact	that had a wooden handle and could breathe well underwater	Transfer	Biology
6. stop sign/mailbox	human-made artifact	that was covered with rust and was crying because it was sad	Transfer	Psychology

Table 2. Recall Prompting Questions.

Recall Prompting Questions
1. What happened at the beginning of the story?
2. What did the kids see or find when they were exploring inside their house?
3. What did the kids see or find when they went outdoors into their backyard?
4. What did the kids see or find along the street outside their house?
5. What did the kids see or find at their neighbor's house?
6. What did the kids see or find in their neighbor's backyard?
7. What did the kids see or find on their way home?
8. What happened at the end of the story?
9. Can you remember anything else about the whole story, from the beginning to the end?

Table 3. Experiment 2 Target Items.

Target Noun Pairs	Ontological Category	Counterintuitive (2 violations) Target Items	Violation Type
1. lizard/rat	animal	that always melted in the hot sun and could never die no matter how old it was	Physics + Biology
2. flower/coconut tree	plant	that would blink at least five times every minute and could disappear and reappear in a different spot in the garden	Biology + Physics
3. banana/mango	plant	that felt angry when it rained and turned invisible every few minutes	Psychology + Physics
4. icicle/rock	natural object	that could give birth to a teapot and liked to sing very loudly	Biology + Psychology
5. hammer/rake	human-made artifact	that could see into the future and could breathe well underwater	Psychology + Biology
6. stop sign/mailbox	human-made artifact	that was floating in midair and was crying because it was sad	Physics + Psychology

Table 4. Experiment 3 Target Items.

Target Noun Pairs	Ontological Category	Target Item Type	
		Intuitive	
1. lizard/rat	animal	that ate insects off of the ground, crawled around quickly on all four of its small feet, and had a long, thin tail	
2. flower/coconut tree	plant	that had roots that went deep into the soil and needed lots of sunshine and water from the rain to grow, and moved from side to side whenever a strong wind blew past it	
3. banana/mango	plant	had a bright yellow skin, smelled very fruity and delicious, and had a fleshy inside that was fresh and ripe	
4. icicle/rock	natural object	that was both very thick and hard, looked shiny in the sunlight, and had a smooth surface that felt quite cold to the touch	
5. hammer/rake	human-made artifact	that felt very heavy to hold, was brown in color, and had a wooden handle made of oak	
6. stop sign/mailbox	human-made artifact	that was made of metal, had sharp edges along its corners, and was covered in a layer of thick red colored rust	
		Counterintuitive (3 violations)	**Violation Type**
1. lizard/rat	animal	that always melted in the hot sun, could never die no matter how old it was, and could hear other creatures' thoughts	Physics + Biology + Psychology
2. flower/coconut tree	plant	that would blink at least five times every minute, could disappear and reappear in a different spot in the garden, and knew everything that had ever happened in the history of the world	Biology + Physics + Psychology
3. banana/mango	plant	that felt angry when it rained, turned invisible every few minutes, and could live in outer space without needing any oxygen	Psychology + Physics + Biology
4. icicle/rock	natural object	that could give birth to a teapot, liked to sing very loudly, and could be in two different places at the exact same time	Biology + Psychology + Physics
5. hammer/rake	human-made artifact	that could see into the future, could breathe well underwater, and could travel back and forth in time	Psychology + Biology + Physics
6. stop sign/mailbox	human-made artifact	that was floating in midair, was crying because it was sad, and ate fire every morning to get energy for the day	Physics + Psychology + Biology