Relational Categories Are More Mutable than Entity Categories

Jennifer Asmuth
Susquehanna University
Dedre Gentner
Northwestern University

Jennifer Asmuth, Department of Psychology, Susquehanna University; Dedre Gentner, Department of Psychology, Northwestern University.

This research was supported in part by a grant from the Office of Naval Research, award number N00014-02-10078, to the second author.

Correspondence concerning this article should be addressed to Jennifer Asmuth, Department of Psychology, Susquehanna University, 514 University Avenue, Selinsgrove, PA 17870.

Email: asmuth@susqu.edu
Relational Categories Are More Mutable than Entity Categories

(Title page stripped of author information)
Abstract
Across three experiments, we explore differences between relational categories—whose members share common relational patterns—and entity categories, whose members share common intrinsic properties. Specifically, we test the claim that relational concepts are more semantically mutable in context, and therefore less stable in memory, than entity concepts. We compared memory for entity nouns and relational nouns, tested either in the same context as at encoding or in a different context. We found that (1) participants show better recognition accuracy for entity nouns than for relational nouns and (2) recognition of relational nouns is more impaired by a change in context than is recognition of entity nouns. We replicated these findings even when controlling for factors highly correlated with relationality, such as abstractness-concreteness. This suggests that the contextual mutability of relational concepts is due to the core semantic property of conveying relational structure and not simply to accompanying characteristics such as abstractness. We note parallels with the distinction between nouns and verbs and suggest implications for lexical and conceptual structure. Finally, we relate these patterns to proposals that a deep distinction exists between words with an essentially referential function and those with a predicate function.

Keywords: relational categories, relational nouns, nouns-verb differences, memory, semantics
Relational Categories Are More Mutable than Entity Categories

Concepts and categories are central to human thinking. Although earlier accounts of categorization focused largely on hierarchically structured taxonomic categories (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Rosch & Mervis, 1975), subsequent research has made it clear that there are distinct kinds of categories. Barsalou’s (1983) explorations of goal-driven categories showed that people form and use categories that subserve particular goal-related activities. Some of these are generated ad hoc—for example, things to take out of your house in case of fire—but others, such as things to eat on a diet, are relatively stable. Ross and Murphy (1999) found that people’s knowledge about foods was organized into taxonomic categories such as fruits and meats, but also into cross-cutting categories such as breakfast foods and desserts. Cross-cultural comparisons have found variation in category type and usage both within and across cultures (Medin, Lynch, Coley & Atran, 1997; Medin, Lynch, & Solomon, 2000). Tyler and colleagues draw on findings from neuroscience to propose that biological and artifact categories differ in their distribution of perceptual versus functional features and in the kinds of correlations between function and perceptual features (Moss, Tyler & Taylor, 2007; Tyler, Moss, Durrant-Peatfield, & Levy, 2000).

In this paper we explore the distinction between relational categories and entity categories. Relational categories are those whose membership is determined by common relational structure (such as extrinsic relations to other entities), rather than by common intrinsic properties (Gentner & Kurtz, 2005; Markman & Stilwell, 2001). For example, for X to be a carnivore, X must eat animals; for X to be a bridge, X must connect two other points or entities. Relational categories contrast with entity categories like tiger, whose members share many intrinsic properties. (Even superordinate entity categories like mammal share common intrinsic
properties, though not as many as for basic-level categories). In contrast, members of the relational category carnivore (e.g., sharks, eagles, tigers, spiders, and Venus flytraps) share virtually no properties—rather, they share the relation that they all eat animals.

This distinction is related to the distinction between actions and objects (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011), but it is not equivalent to it. While many actions are relations (e.g., x chases y; x gives y the z), others are not (x sneezes; x sleeps). In the other direction, the relations around which category members cohere may involve actions (as in ‘X eats animals’), but they need not (as in passenger or barrier). The distinction between relations and properties is also related to the distinction between functional and perceptual information (Tyler et al., 2000), but again it is not reducible to this distinction—first, because relational information is a broad class that includes functional information as a subclass; and, second, because intrinsic properties need not be perceptual: for example, warm blood is an intrinsic property of mammals and birds, even though it is normally perceptually unobservable.

In this research, we contrast relational categories with entity categories, focusing on how they are encoded and remembered when presented in semantic combination. However, before going on, we need to discuss the key prerequisite: how to identify relational categories. In some cases, morphology provides a clue; relational nouns are often derived from Latinate verbs in a morphologically transparent way, as in destroy \(\rightarrow\) destruction or repent \(\rightarrow\) repentance. However, this is only a partial guide, which misses many relational nouns (e.g., friend, enemy, mother). A second, more general criterion for classifying a noun as referring to a relational category is whether it takes other nouns as arguments (Barker & Dowty, 1993). For example, to be a friend, you have to be a friend of someone, but you can be a person without any dependence on external entities. Likewise, the relational category barrier implies a scenario with three arguments: a
RELATIONAL CATEGORIES ARE MORE MUTABLE

figure, something that blocks access, and a goal, as in “The ranches form a barrier to the migration of the great herds towards their summer pastures.” Although this is a useful criterion, it too misses some cases. Often, not all the arguments of a relational category are obligatory. For example, we can say “A friend in need is a friend indeed” without specifying the role filler for friend of. Moreover, some relational categories (e.g., carnivore) don’t need arguments at all. Therefore, we propose a third test for relational categories: a test we have dubbed the ‘Fetch Test.’ If you were seeking a member of the category, could you identify it just by its own properties or would you have to consider its interactions/relations with other things? For an entity category like tiger, you need only locate the entity itself—it’s intrinsic properties suffice to identify it. But for a relational category like carnivore, intrinsic properties are not enough: you need to verify its relations to other entities (namely, that it eats animals). While you know which items in the room are bananas just by looking at them, you can’t tell which things are gifts without knowing about their relations to other entities. It is more challenging to apply the Fetch Test to abstract nouns, and at the extreme it may become unusable. However, even for fairly abstract nouns like object (an abstract entity noun) or consequence (an abstract relational noun), the general principle will still hold: can you categorize an object based solely on its own properties or must you investigate its relationships to other entities (an object is a bounded material entity, a consequence is something that follows from a set of conditions).

The relative obscurity of relational categories in accounts of categorization might suggest that these kinds of categories are rare in number, or that they are rarely used. But this is not the case; relational nouns figure prominently in adult discourse. In fact, our informal ratings of the 100 highest frequency nouns in the British National Corpus suggest that close to half are
RELATIONAL CATEGORIES ARE MORE MUTABLE

 relational. Furthermore, many superordinate categories (e.g., carnivore, barrier) are relational, even if their members (e.g., shark, wolf, praying mantis; fence; chasm; river) are entity nouns.

Until recently, nominal relational categories have been largely ignored. Most category research has focused on entity categories—categories like cat or apple that can be characterized in terms of shared intrinsic features among members. One of the earliest explorations was that of Barr and Caplan (1987), who contrasted categories defined by intrinsic features with those defined by extrinsic features and found differences in judgments of typicality and membership between the two types. Since then, the idea that relational categories are an important class of human conceptual groupings has gained ground, and researchers have begun to characterize their properties (Feist & Cifuentes-Ferez, 2007; Gentner, 2005; Gentner & Kurtz, 2005; Goldwater & Markman, 2011; Markman & Stilwell, 2001; Rehder & Ross, 2001).

As with other kinds of categories, relational categories highlight similarities (in this case, relational similarity) among their members. For example, when objects are presented as belonging to the same relational category (e.g., polar bear and seal as predator) their perceived similarity increases compared to that of the same pair presented as members of different relational categories (e.g., polar bear as predator, seal as prey) (Jones & Love, 2007). Another parallel with standard categories is that comparing exemplars aids in learning relational categories (Christie & Gentner, 2010; Gentner, Anggoro, & Klibanoff, 2011; Kotovsky & Gentner, 1996; Kurtz, Boukrina, & Gentner, 2013; Goldwater, Bainbridge, & Murphy, 2016) just as it does for entity categories (Higgins & Ross, 2011). Further, receiving a relational category label aids children in identifying new members of the category (Christie & Gentner, 2014; Gentner et al., 2011).
There are also differences between relational categories and entity categories. Goldwater, Markman and colleagues found that role-governed categories (e.g., thief and victim, both roles in a robbery schema) are likely to be described by ideal features, whereas entity categories are likely to be described by typical features (Goldwater, Markman, & Stilwell, 2011). Further, in category-based induction, the most effective standard for relational role categories is an ideal member (Rein, Goldwater & Markman, 2010), rather than a typical member as for entity categories. Relatedly, Kloos and Sloutsky (2008) make a distinction between natural kind concepts like bird (roughly corresponding to what we are calling entity categories) and nominal kind concepts based on rules (which partly overlap with relational categories). Using a category learning task, they found that people learned natural kind categories better by observing examples than by being given an explicit rule characterizing membership criteria. The reverse was true for nominal categories.

Verbs and relational categories. A useful analogy can be made between relational nouns and verbs. Relational terms like verbs and prepositions differ from concrete noun categories not only in their grammatical privileges, but in their semantics (Croft, 1991, 2001; Gentner, 1981, 1982; Gentner & Boroditsky, 2001; Graesser & Hopkinson, 1987; Huttenlocher & Lui, 1979; Kersten, 1998; Kersten & Earles, 2004; Langacker, 1987, 2008; Pavlicic & Markman, 1997; Vigliocco et al., 2011). We suggest that some of the same contrasts apply to relational nouns vs. entity nouns. As noted earlier, many relational nouns are derived from verbs (e.g., robbery and rob). More to the point, they often retain the argument structure associated with the verb. For example, if we know that Jesse James robbed the Glendale train of its gold, then we can speak of the robbery of the gold from the train by Jesse James. Thus, relational nouns resemble verbs in
that they often take arguments (Barker & Dowty, 1993). (This is sometimes referred to as being *semantically unsaturated* (e.g., de Bruin & Scha, 1988).

Beyond these derivational relationships, there are deeper parallels between relational nouns and standard relational terms like verbs and prepositions, stemming from the fact that their meanings center around extrinsic relations with other concepts. Both semantically and syntactically, the behavior of relational nouns resembles that of verbs.

Based on these parallels, Gentner and Kurtz (2005) proposed the following analogy as a framework for research:

\[
\text{Relational Nouns} : \text{Entity Nouns} : : \text{Verbs} : \text{Nouns}
\]

This analogy suggests that findings on processing distinctions between nouns and verbs can be used to generate parallel hypotheses concerning relational nouns and entity nouns. Gentner (1981, 1982, 2005; Gentner & Kurtz, 2005) laid out a set of phenomena that differentiate verbs from concrete nouns: verbs are slower to be acquired (Bornstein et al., 2004; Caselli et al., 1995; Gentner, 1982; Gentner & Boroditsky, 2001, 2009); harder to remember, (e.g., Kersten & Earles, 2004); more mutable in meaning under semantic strain (Gentner & France, 1988); and more polysemous (controlling for word frequency). Many of these contrasts also apply to relational nouns as compared to entity nouns. For example, relational nouns are acquired later than entity nouns (Gentner, 2005; Gentner & Rattermann, 1991; Hall, 1994), just as verbs are acquired later than nouns. Even when relational nouns enter the vocabulary, their full meanings are often not understood until rather late, in contrast with children’s early grasp of the meanings of many entity nouns (Gentner, 2005). Indeed, children often seem to mistake relational nouns for object-reference terms (e.g., *uncle = friendly man with a pipe*) before evolving to the relational
interpretation held by adults (e.g., *brother of mother or father*) (Clark, 1973; Gentner, 2005; Gentner & Rattermann, 1991; Keil & Batterman, 1984; Hall & Waxman, 1993).

In this paper we focus on two interrelated differences between nouns and verbs, and ask whether they might also apply to entity nouns vs. relational nouns: (1) verbs are harder to remember than nouns, in both recognition and recall (Kersten & Earles, 2004; Earles & Kersten, 2000); and (2) verbs are more mutable than nouns—that is, more context-sensitive in their semantic interpretation (Gentner, 1981; Gentner & France, 1988; Reyna, 1980). By *mutability* we mean a word’s propensity to take on conceptually different encodings in different contexts. Gentner and France (1988) compared nouns and verbs along this dimension by asking participants to paraphrase semantically strained sentences such as *The lizard worshipped.* The results showed greater semantic change for the verbs than for the nouns; for example, one participant paraphrased the above sentence as “*The small grey reptile lay on a rock and stared unblinkingly at the sun.*” As this example suggests, participants appeared to take the noun as referring to a member of an established category (e.g., lizard = small gray reptile) and to interpret the verb in terms of that referent: e.g., as ‘some act that a lizard could perform that could be described as praying.’

That verbs are less well recalled and recognized than nouns is related to their mutability. To the extent that a word’s meaning adapts to its encoding context (mutability), we would expect that word to fare poorly in memory, especially if the context changes (Gentner, 1981; Kersten & Earles, 2004). Studies by Kersten and Earles and colleagues have borne out this prediction (Earles, Kersten, Turner, & McMullen, 1999; Earles & Kersten, 2000; Kersten & Earles, 2004). For example, Kersten and Earles (2004) tested recognition memory for simple noun-verb sentences and found (a) better recognition for nouns than for verbs and (b) a greater drop in
accuracy for verbs than for nouns when the accompanying contextual words were changed from encoding to recognition. This fits with the mutability claim: because verb encodings are adapted to fit their noun arguments, they are recognized less well than nouns, especially if the context changes.

If the above analogy holds, then we should find that this pattern—greater contextual mutability with concomitant lowering of recognition accuracy—holds between relational nouns vs. entity nouns, just as for verbs vs. nouns. If so, then when relational nouns are combined with entity nouns, the relational nouns should be both less stable under paraphrase and less well retained in a memory task.

To test these claims, in Experiment 1, we presented participants with noun-noun compounds, each containing an entity noun and a relational noun. Both entity noun-relational noun (ER) order and relational noun-entity noun (RE) order were used (e.g., tooth opponent vs opponent tooth), between subjects. Properties of noun–noun conceptual combinations have been well-studied and provide an appropriate arena in which to study the mutability of concepts (Costello & Keane, 2000; Hampton, 1987; Murphy, 1988; Wisniewski, 1996, 1998). Adopting the logic of the Kersten and Earles studies, we then compared recognition for the entity nouns versus the relational nouns, given either the old context word or a new context word. If, as we hypothesize, the encoding of relational nouns is highly context-sensitive, then during study the relational nouns should adapt their meaning to their entity partners; thus they are likely to fare poorly in the later surprise recognition test. In contrast, the encoding of entity nouns is predicted to be relatively context-independent; thus, entity nouns should retain their intrinsic character from the study situation to the test situation (regardless of their relational partners). Therefore, at
the recognition test, people should be more sensitive to a change in the entity noun than to a change in the relational noun.

**Experiment 1**

In our first study, we employed a recognition paradigm based on Kersten and Earles’ (2004) investigation of noun and verb recognition. Participants were given a list of conceptual combinations consisting of an entity noun (E) and relational noun (R) (e.g., *a leg reference*) and asked to rate how difficult the phrases were to understand on a scale from 1 to 7. This task served to encourage participants to actively comprehend the sentences, as well as providing a means of calibrating the difficulty of the items.

After reading and rating the phrases, participants were given an unrelated filler task, followed by a surprise recognition test that included the old conceptual combinations (E \(_{\text{old}}\)R \(_{\text{old}}\)) seen in the ratings task, combinations of old entity nouns and new relational nouns (E \(_{\text{old}}\)R \(_{\text{new}}\)), combinations of old relational nouns and new entity nouns (E \(_{\text{new}}\)R \(_{\text{old}}\)), and conceptual combinations composed solely of novel words (E \(_{\text{new}}\)R \(_{\text{new}}\)). The task was to say whether each whole phrase had been seen in the original ratings task. For example, if a participant saw *a truck limitation* during the ratings task, she would see one of the phrases in Table 1 during the recognition test. (The first phrase is the original phrase seen at study; the others are new combinations.)

Old nouns were always combined with new nouns except when they appeared with their original partner at encoding. That is, if both *truck limitation* and *book threat* were seen at encoding, then a new combination of two “old” nouns like *truck threat* would not be seen at test. This allowed us to distinguish between the false alarms triggered by the E \(_{\text{old}}\) and those triggered by R \(_{\text{old}}\). We expected high hit rates for the E \(_{\text{old}}\)R \(_{\text{old}}\) combinations, since these phrases were
actually seen by participants. Likewise, an E\textsubscript{new}R\textsubscript{new} combination should elicit a very low false alarm rate since neither word had been seen at study.

The logic of the study is as follows. Because relational nouns are more context-sensitive than entity nouns, the meanings of the relational nouns should be more dependent upon their context and therefore more susceptible to a shift in meaning between different contexts. For example, suppose a participant instantiates the phrase *a truck limitation* at encoding as “a vehicle that is hard to parallel park.” If, at recognition, the same participant sees the E\textsubscript{old}R\textsubscript{new} phrase *a truck threat* and interprets it as “a runaway vehicle,” she should feel a sense of overlap based on the common concept of *vehicle*; thus she may be likely to false-alarm to the phrase. However, if she sees the E\textsubscript{new}R\textsubscript{old} combination *a book limitation* at test, she might think of “a boring story.” This should be comparatively less similar to her prior encoding of “hard-to-park vehicle” and therefore less likely to trigger a false alarm. Thus, the prediction is that people will show greater recognition accuracy (and therefore fewer false alarms) to phrases of new entity nouns and old relational nouns (E\textsubscript{new}R\textsubscript{old}) than to phrases with old entity nouns and new relational nouns (E\textsubscript{new}R\textsubscript{old}).

**Method**

**Design.**

For stimuli presentation, we used a 2x2x2 mixed-factorial design. Phrase word order—the order of the entity and relational noun within the noun-noun phrase (ER = entity noun followed by relational noun; RE = relational noun followed by entity noun)—was manipulated
between participants. The variables of entity noun status (whether the entity noun in the phrase was old or new) and relational noun status (whether the relational noun in the phrase was old or new)—were manipulated within participants. The factors for stimuli presentation were used to derive a different design for hypothesis testing, in which we compare false alarms in a 2 (word order) x 3 (recognition phrase) design, across both participants and item, as described in the Results section.

Participants

Fifty-three undergraduates participated in this study in exchange for partial course credit. While we did not track gender and age for this set of studies, about 65% of undergraduates participating in the lab’s studies tend to be female. All participants are fluent in English, with 83% speaking only English and an additional 10% reporting both English and another language as their first language.

Materials

The encoding materials consisted of 96 conceptual combination phrases: 64 novel entity noun–relational noun combinations, and 32 highly conventional noun-noun combinations (e.g., *noodle casserole* or *football stadium*), which served as filler items.

The recognition materials consisted of 32 conceptual combinations: 8 phrases from the original list at study (E\text{old}R\text{old} or R\text{old}E\text{old}), 8 phrases with an old entity noun and new relational noun (E\text{old}R\text{new} or R\text{new}E\text{old}), 8 phrases with a new entity noun and an old relational noun (E\text{new}R\text{old} or R\text{old}E\text{new}), and 8 phrases with a new entity and new relational noun (E\text{new}R\text{new} or R\text{new}E\text{new}). These were always presented in the same word order (ER or RE) as seen at study. One of the E\text{new}R\text{new} phrases and one of the E\text{new}R\text{old} phrases was omitted from analysis due to experimenter error.
The nouns were matched in frequency according to the norms of Kucera and Francis (1967). All nouns fell in the frequency range of 20-100. The average entity noun frequency was 48.43; average relational noun frequency was 47.00. There was no reliable difference in frequency between the two lists ($t < 1$).

Stimuli were initially selected by the experimenters, who judged whether a referent named by the noun could be identified primarily by intrinsic properties (entity nouns) or required knowledge of extrinsic relationships to other entities (relational nouns). Undergraduate informants who were blind to predictions of the study provided additional categorization judgments and nouns for which there was substantial agreement were selected as stimuli. Furthermore, a subset of these stimuli (68%) were later included in the ratings task to select nouns for Experiment 3 (see Supplementary Materials for instructions). An interrater reliability analysis comparing the categorization by these expert judges to those used in Experiments 1 and 2 yielded substantial agreement, kappa = .70.

**Procedure**

The participants’ task was to interpret each phrase and to rate the difficulty of doing so. To encourage naturalistic encoding, we asked participants to imagine that they had overheard the phrase in passing, and to consider what the speaker would most likely have meant; then they were to rate the difficulty of constructing the interpretation on a scale from 1 to 7 (see Supplementary Materials for instructions). The ratings task served to keep participants on-task and actively interpreting the phrases. We did not ask participants to write out the paraphrases, because we feared that writing 96 paraphrases would be too onerous and time-consuming. However, in a pilot study we obtained paraphrases from a highly overlapping set of materials (that were also used in Experiment 2). As described below, the results of that task showed that
participants understood the instructions and were able to construct sensible interpretations of the conceptual combinations, in both ER order and RE order. For example, the phrase *insect goal* in ER order was paraphrased as *a basic, small future achievement* and *the objective of a colony of bugs*, while in RE order *goal insect* elicited *bug you are looking for* and *a small bug that always has a mission* (see Supplementary Materials for additional example paraphrases and difficulty ratings).

Participants completed the paper-based ratings task for the 96 conceptual combinations, participated in an unrelated filler task for 20 minutes, and then completed the paper-based surprise recognition task of 32 conceptual combinations. We did not ask participants to make difficulty judgments during the surprise recognition test out of concern that such processing would affect the recognition task itself.

**Results and Discussion**

To preview, the results conform to predictions: semantic context had a greater effect on memory for relational nouns than on memory for entity nouns. We begin by reporting overall accuracy of participant responses, then discuss false alarms to test phrases across participants, before moving to items analyses.

As expected, participants were highly accurate in correctly endorsing old noun–noun combinations as previously seen \( (M = .83, SD=.18 \text{ across both orders}) \) and in rejecting entirely new combinations (false alarm rate \( M = .10, SD=.15 \text{ across both orders} \)). Clearly, participants had been attending to the task. We argue that relational nouns are more poorly remembered than entity nouns and more influenced by their context. Therefore, at test participants are more likely to incorrectly identify phrases consisting of new relational nouns paired with older entity nouns as previously seen than phrases with new entity nouns with old relational nouns.
This prediction is tested most directly by examining false alarms to the three kinds of phrases that include novel nouns at test via a 2 (word order: ER/RE) x 3 (recognition phrase: \(E_{\text{new}}R_{\text{new}}, E_{\text{new}}R_{\text{old}}, E_{\text{old}}R_{\text{new}}\)) ANOVA across participants. The analysis revealed no main effect of word order, \(F(1, 51) = 1.49, p = .23, \eta^2_p = .03\). However, it’s worth noting that study phrases in the RE order were rated as more difficult on the 1-7 scale (\(M = 5.21, SD = .71\)) than those in the ER order (\(M = 4.80, SD = .92\)), \(t(63) = 3.60, p = .001\) via a paired samples t-test across items. We return to this issue later in the discussion.

[Please place Table 2 about here]

The participant analysis revealed a main effect of the within-subjects variable, recognition phrase, \(F(2, 102) = 20.69, p < .001, \eta^2_p = .29\). Planned comparisons revealed that participants produced more false alarms to \(E_{\text{old}}R_{\text{new}}\) phrases (\(M = .30, SD = .24\)) than both \(E_{\text{new}}R_{\text{old}}\) phrases (\(M = .19, SD = .19\), \(p = .001\), and \(E_{\text{new}}R_{\text{new}}\) phrases (\(M = .10, SD = .15\), \(p < .001\). Participants experienced the least sense of familiarity with phrases consisting of two new words, and these had the lowest false alarms. In contrast, they produced false alarms at a higher rate to phrases that contained a noun they had seen previously and they produced more false alarms to phrases with old relational nouns than with old entity nouns.

These results are consistent with the idea that the entity nouns were encoded in a relatively context-independent manner during study (as well as during test); thus, entity nouns retained their intrinsic referential character across these two situations, regardless of their relational partner, leading to the sense that the phrase had been encountered previously and eliciting false alarms. For example, in the pilot study, the conceptual combination *tooth*
opponent was paraphrased as cavity. Apparently the relational noun (opponent) was interpreted as “something that opposes a tooth”—hence, a cavity. On this account, an old relational noun paired with a new entity noun would not trigger a sense of familiarity, because the entity noun is novel and the relational noun would receive a different interpretation in its new context.

We also examined false alarm rate across items. Because word frequency may influence retrieval (for example, low frequency—and therefore distinctive—words fare better at recognition), we controlled for word frequency (Kucera & Francis, 1967) in the development of our stimuli and use frequency as a covariate in the items analysis. We also employ word length as a covariate, since longer words tend to be more easily recognized than shorter words. Because approximately 30% of the test stimuli in Experiments 1 and 2 had no recorded concreteness value in the MRC database, we were not able to use concreteness as a covariate. We address the limitations of this issue on our interpretations in greater depth in the discussion of Experiment 2.

An analysis of false alarms across items via a 2 (word order: ER/RE) x 3 (phrase type: E_{new}R_{new}, E_{new}R_{old}, E_{old}R_{new}) mixed-factorial ANCOVA showed a similar pattern to that of the participant analysis. There was no main effect of word order within the phrase, $F_{2}(1, 15) = .22, p = .64, \eta^2_p = .02$. However, there was a significant effect of phrase type on false alarms across both ER and RE orders, after controlling for the effect of entity and relational noun frequencies (Kucera & Francis, 1967) and word length, $F_{2}(2, 15) = 3.98, p = .04, \eta^2_p = .35$. A planned comparison revealed that phrases with a new relational noun had the highest false alarm rate at the recognition test ($M = .28, SD = .10$), differing significantly from phrases with a new entity noun ($M = .17, SD = .09$), $p = .046$, and phrases with both new entity and new relational nouns ($M = .13, SD = .10$), $p = .02$. The covariate word length of relational noun was also significantly
RELATIONAL CATEGORIES ARE MORE MUTABLE

related to false alarms to the test phrases\(^1\), \(F(1, 15) = 12.48, p = .003, \eta^2_p = .45\). The items analyses also suggest that participants largely failed to detect changes in the relational nouns.

The results of this study bore out our predictions: participants experienced a greater sense of (incorrect) familiarity to phrases with an old entity noun and a new relational noun than phrases with an old relational and a new entity noun. This provides encouragement for our hypothesis that relational nouns are more mutable than entity nouns. However, the study design, in which people are asked to recognize the whole phrase, makes it difficult to isolate the effects of entity vs. relational nouns. The results are consistent with our story that entity nouns are more stable across contexts, and relational nouns more mutable. But it is also possible that relational nouns are simply harder to recognize in general, regardless of context. Thus in Expt. 2, participants were again given the same four kinds of phrases, but were asked to recognize just one of the terms—entity or relational noun.

A second alteration made in Experiment 2 concerns word order. Although there was no main effect of word order on recognition, phrases in the RE order (e.g., *opponent tooth*) were judged more difficult during the interpretation task at study than the same combinations in ER order (e.g., *tooth opponent*). This difference is consistent with discussions of word order in conceptual combination (Costello & Keane, 1996, 2000; Gerrig & Murphy, 1992; Murphy, 1988; Wisniewski, 1996, 1998). In noun-noun combinations, the second noun serves as the head noun and the first noun serves as modifier. Conceptual combinations are often understood by using the modifier to fill a slot in the head noun schema (Murphy, 1990; Wisniewski & Gentner, 1991; Wisniewski, 2001). Thus it makes sense that a conceptual combination is most natural when the relational noun occupies the head position, where its well-established relational structure provides slots that can be filled by the entity noun in the modifier position. Thus, in the
combination *tooth opponent, opponent* would be taken as something that blocks teeth from reaching some goal. Of course, *opponent tooth* can also be interpreted (for example) as an adversarial tooth; but this order may be more difficult to develop.

Although the major results of Experiment 1 did not appear to be affected by word order, we were concerned that the perceived difficulty might compromise the naturalness of the interpretations. Therefore, we conducted a pilot study to test whether people could construct reasonable interpretations in either order. We asked 21 participants to paraphrase and provide difficulty ratings for the 32 conceptual combinations used at test in Experiment 2 (largely overlapping with those of Experiment 1). Each participant received 16 paraphrases in RE order and 16 in ER order, counterbalanced across participants.

Participants generally managed to generate sensible paraphrases for either order (e.g., *an insect goal* was paraphrased as *a small pursuit*, while *a goal insect* was *a bug you are looking for*; see Supplementary Materials for sample paraphrases and difficulty ratings). However, they found RE order more difficult (*M* = 4.42, *SD* = .71) than ER order (*M* = 4.10, *SD* = .86), *t*(20) = 2.45, *p* = .02, by a two-tailed, paired samples t-test over subjects. Items analysis also suggests that RE order was more difficult (*M* = 4.42, *SD* = 0.83) than ER order (*M* = 4.11, *SD* = 1.06) and this approached significance via a 2-tailed, paired samples t-test, *t*(31) = 1.68, *p* = .10. These results buttress the concern that participants find the RE order more challenging (although this difference between word orders could also reflect a selection bias in the stimuli, since the words were not randomly chosen). Therefore, in the remaining studies, we used only the ER order.

However, using the ER order has the additional advantage of biasing somewhat against our prediction that the encodings of the relational nouns will be adapted to fit the meanings of the entity nouns. This is because in noun-noun combinations, the second noun (often called the
head noun) is generally taken to be the referent, with the first noun acting as a modifier or
predicate applying to the head noun (Cohen & Murphy, 1984; Murphy, 1988, 1990; Wisniewski,
1996, 1997). For example, a book tiger might be a tiger who likes to read, whereas a tiger book
is a book about tigers. An acrobat hippopotamus is a graceful hippopotamus, whereas a
hippopotamus acrobat is probably a very clumsy acrobat (Gentner & France, 1988). In the ER
order, the relational noun is in the head noun position, where the usual processes of noun-noun
combination would render it more stable, not less as we predict.

Experiment 2

In order to isolate the encoding mutability effect, in Experiment 2 we adapted Kersten
and Earles’ (2004) noun–verb recognition paradigm to test recognition for single words within
noun–noun phrases. As in Experiment 1, participants first read phrases consisting of an entity
noun and a relational noun and were asked to judge the difficulty of constructing a meaningful
paraphrase of the conceptual combination. At recognition, they saw phrases with one word
indicated as the target word to be recognized. The target word was paired with either an old or a
new context word.

If the encoding of relational nouns is more influenced by context than that of entity
nouns, then entity nouns should be more stable across contexts. This means that recognition for
relational nouns should be more impaired by a new context than that of entity nouns. Using d-
prime as the measure of recognition accuracy, we predict that (1) d-prime should be higher
overall for entity nouns than for relational nouns; (2) d-prime should be lower in the new context
than in the old; and (3) there should be an interaction between noun focus (entity or relational)
and recognition context (old or new): that is, d-prime should drop more over a change in context
for relational nouns than for entity nouns.
Method

Design

We manipulated three independent variables to create the stimuli: noun focus (whether the participant was tested for entity nouns or relational nouns at recognition), target status (whether the tested noun was old or new), and recognition context (whether the target noun at test was paired with an old or new noun at test). Noun focus was manipulated between participants; target status and recognition context were manipulated within participants. These factors were used to generate different dependent variables for hypothesis testing, as described below. For the subjects analysis, the dependent measures were hits to old targets and false alarms to new targets; these were used to compute d-prime. For the items analysis, the dependent measure was percent correct.

Participants

Seventy-two undergraduates from the same university participant pool as those in Experiment 1 participated in exchange for partial course credit.

Materials

The encoding materials were drawn from our stimulus list for Experiment 1 (with corrections made for two phrases that were incorrectly combined). They consisted of 96 conceptual combination phrases in ER order: 64 relational noun–entity noun combinations and 32 conventional combinations (e.g., noodle casserole or football stadium) as filler items.

Procedure

Paralleling the procedure in Experiment 1, participants completed the computer-based ratings task for the 96 conceptual combinations. They were instructed to consider what the
phrase might mean in a natural context, such as overhearing the phrase in a college dining hall, and to rate the difficulty of interpretation on a scale from 1 to 7 (see Supplementary Materials for instructions). After a 20-minute unrelated filler task, participants completed the computer-based surprise recognition task. On the recognition test, participants saw 32 noun–noun conceptual combinations (eight each of E$_{\text{old}}$R$_{\text{old}}$, E$_{\text{old}}$R$_{\text{new}}$, E$_{\text{new}}$R$_{\text{old}}$, and E$_{\text{new}}$R$_{\text{new}}$ phrases) and judged whether the target word (displayed in red) had been seen during the earlier ratings task or was new.

**Results**

The results of the recognition test were analyzed using a 2 (noun focus: entity or relational, between-subjects) x 2 (context: old or new, within-subjects) design. Following MacMillan and Kaplan (1995), we computed recognition accuracy ($d'$-prime) for both old contexts and new contexts for each participant. $D'$-prime is a signal detection statistic that measures an individual’s ability to discriminate between different stimuli (e.g., between previously seen and new items), taking individual bias into account. One $d'$-prime measure assessed a participant’s ability to distinguish old (previously seen) nouns from new nouns when they were seen in an ‘old’ context—that is, in the same context as at encoding. For this old context measure, we calculated (a) the proportion of hits to old nouns presented at test with the same old context noun as at encoding ($E_{\text{old}}R_{\text{old}}$) and (b) the proportion of false alarms to new nouns presented with old context nouns (that is, $E_{\text{new}}R_{\text{old}}$ for participants with an entity noun focus and $E_{\text{old}}R_{\text{new}}$ for those with a relational noun focus; see Table 3.) In this and subsequent studies, for participants in which mean hits were 100% or false alarms were 0%, we adjusted $d'$-prime to avoid infinite values by using the formulas of MacMillan and Creelman (2005). Hits to $E_{\text{old}}R_{\text{old}}$ phrases were very similar for both entity noun focus (.73) and relational noun focus
RELATIONAL CATEGORIES ARE MORE MUTABLE

(.74). However, the relational noun focus elicited more false alarms (.39) to new nouns presented with old context nouns than the entity noun focus (.27).

[Please place Table 3 about here]

Similarly, to assess participants’ recognition accuracy in new contexts, we calculated (a) the proportion of hits to old nouns presented with a new context noun at test ($E_{\text{old}R_{\text{new}}}$ for participants with an entity noun focus; $E_{\text{new}R_{\text{old}}}$ for those with a relational noun focus) and (b) the proportion of false alarms to new nouns presented with a new context noun ($E_{\text{new}R_{\text{new}}}$). Hits and false alarms were .59 and .17, respectively, for the entity noun focus condition and .50 and .28 for the relational noun focus condition. These proportions were used to compute d-prime for a new context, as shown along with d-prime for the old context in Figure 1.

As predicted, a mixed 2 (noun focus: entity, relational) x 2 (context: old, new) ANOVA across participants revealed that participants tested on entity nouns showed better recognition accuracy ($M = 1.34, SD = 0.75$) than those tested on relational nouns ($M = .87, SD = 0.70$) across both old and new contexts, $F(1, 70) = 13.40, p < .001, \eta^2_p = .16$). They also showed marginally better recognition accuracy for nouns in an old context ($M = 1.20, SD = 0.69$) than in a new context ($M = .98, SD = 0.82$), $F(1, 70) = 3.75, p = .06, \eta^2_p = .05$. With respect to our third prediction, recognition accuracy for relational nouns appeared to decline from old context ($M = 1.05, SD = 0.69$) to new context ($M = 0.69, SD = 0.68$); in contrast, performance for entity nouns did not vary much from old context ($M = 1.37, SD = 0.66$) to new context ($M = 1.32, SD = 0.84$). However, although these findings fit the predicted pattern, the noun focus x context interaction was not significant, $F(1, 70) = 2.01, p = .16, \eta^2_p = .03$. 


We ran an items analysis (ANCOVA) using noun focus (entity, relational), recognition context (old, new), and target status (whether the tested noun was old or new) as independent variables and correct response (endorsing old nouns as old and rejecting new nouns as new) as the dependent variable, with word frequency and word length as covariates. After controlling for the effects of the covariates, we did not find the expected main effect of noun focus across items, $F(2,154) = 1.93, p = .17, \eta^2_p = .04$; an effect of context, $F(2,154) = 1.03, p = .32, \eta^2_p = .02$; or a noun focus x context interaction, $F(2,154) = 1.02, p = .32, \eta^2_p = .02$. Neither covariate was significantly related to correct response.

**Discussion**

These results are consistent with our hypothesis that relational nouns are more mutable in context than entity nouns. At test, people showed greater recognition accuracy for entity nouns than for relational nouns (as well as for nouns in old contexts than those in new contexts). This pattern parallels that for nouns and verbs reported by Kersten and Earles (2004). They found that people were better able to recognize old nouns across different verbs than old verbs across different nouns, and noted that this pattern could result from the greater contextual mutability of verbs relative to nouns (see Gentner, 1981; Gentner & France, 1988). Our findings invite an analogous conclusion for relational nouns as compared to entity nouns.

Our third prediction was that a change in context would have more adverse effects on recognition of relational nouns than on recognition of entity nouns. Although the results were in the predicted direction, the interaction between noun focus and recognition context was not
significant. This led us to ask whether that additional characteristics of the nouns used here—such as abstractness—could have influenced recognition performance.

Abstractness and relationality. In constructing the materials, we noted an intriguing pattern that seems to hold across our materials and in the lexicon more broadly—namely, a strong association between relationality and abstractness. Of course, there are relatively concrete relational words, such as father, and abstract entity words, like idea. However, overall it is very likely that a relational word will also be abstract. For example, words like contribution, advice, and majority all denote important relational categories that have very few attributes available to the senses. Even if we consider relational concepts that have concrete exemplars (such as barrier), we note that the exemplars lack common intrinsic attributes. The glue that holds them together is the abstract idea of “preventing someone (or something) from reaching its goal (or destination).”

This connection between abstractness and relationality raises some fascinating theoretical issues—such as a possible link between our capacity for relational thinking and our capacity for abstract thought (Gentner & Asmuth, in preparation; Gentner, 2003, 2010; Jamrozik, McQuire, Cardillo, & Chatterjee, 2016; Penn, Holyoak & Povinelli, 2008). We return to this issue in the General Discussion. However, on a methodological level, it raises a question concerning the interpretation of the previous two studies. It is well established that imageability and concreteness influence recognition of nouns (Morris & Reid, 1974; Gorman, 1961). Thus, if indeed abstractness and relationality are positively correlated, then the results of our first two studies could be due to high vs. low abstractness (or its reverse, imageability) rather than to differences in relationality. This called for further scrutiny.
When we examined our stimulus set, we found substantial differences in abstractness between the entity nouns and relational nouns. About 2/3 of the stimuli were included within the MRC database (Coltheart, 1981). Comparison of these norms showed that the entity nouns were reliably more concrete ($M = 589.96, SD = 371.10$) than the relational nouns ($M = 386.33, SD = 96.87$), $F(1, 42) = 87.71, p < .001$, and also more imageable ($M = 585.39, SD = 53.01; M = 422.17, SD = 100.64$, respectively), $F(1, 42) = 43.76, p < .001$. These differences are consistent with a deep connection between relationality and abstractness, as discussed below. But with respect to the current studies, they leave open the possibility that differences in abstractness may influence the results—and possibly even account for the greater mutability of relational terms.

To decide this question, we needed to isolate the effect of relationality on mutability. Therefore, in Experiment 3 we designed a set of relational and entity nouns that were matched for concreteness. By equating concreteness, we can discover whether relationality per se affects mutability and recognition accuracy.

In order to focus on the encoding and recognition of each kind of term separately, we used a somewhat different design from the previous two studies. Instead of combining entity and relational nouns in the same encoding phrase, we combined each of the experimental nouns with a neutral “host noun” (e.g., marigold). Thus, each phrase had a host noun and either an entity noun (e.g., marigold palace) or a relational noun (e.g., marigold sister). The experimental nouns were always the head nouns in these phrases, and the host noun was the modifier; these host nouns were always entity nouns (see below). Another alteration from the prior studies is that target noun type (relational or entity noun) was a within-subject factor in this study. In this way we could compare recognition for relational nouns and entity nouns within participants, with encoding context equated across groups.
In the recognition test, participants saw entity and relational nouns paired with either the same host noun as at study or a different host noun. The key predictions are as before: (1) that memory for entity nouns will be better than for relational nouns; (2) that recognition in an old context will be better than recognition in a new context; and (3) that a change in context will impair recognition accuracy for relational nouns more than that for entity nouns.

**Experiment 3**

As in the prior two studies, we followed a two-part plan. First, participants saw a list of conceptual combinations (e.g., marigold sister) consisting of a host noun (H) paired with either an entity noun (E) or a relational noun (R), which they rated for difficulty in interpretation. They later received a surprise recognition test which assessed their ability to discriminate new from old entity nouns or relational nouns in new and old contexts—that is, the target noun paired with either the same host noun as at study or with a different host noun. For example, if a participant saw a phrase like marigold sister during the ratings task, she would see one of the phrases in Table 4 during the recognition test. As in prior studies, participants never saw combinations of two experimental nouns during test: for example, if both marigold sister and albatross citizen were seen at encoding, then marigold citizen would never be seen at test.

We expected high hit rates for target nouns in the original combinations ($H_{\text{old}}R_{\text{old}}$ and $H_{\text{old}}E_{\text{old}}$), since participants will have actually seen the phrase before. Likewise, target nouns in $H_{\text{new}}E_{\text{new}}$ and $H_{\text{new}}R_{\text{new}}$ combinations should elicit a very low false alarm rate since both words are entirely new. As before, the predictions concern recognition accuracy ($d'$-prime) for entity nouns versus relational nouns across contexts: (1) $d'$-prime should be higher overall for entity nouns than for relational nouns; (2) $d'$-prime should be lower in new contexts than in old
contexts; and (3) d-prime for relational nouns will be lower in a new context than will that for entity nouns.

[Please place Table 4 about here]

Method

Design

There were three within-subjects factors: noun focus at recognition test (whether the target was entity or relational), target status (whether the target noun was old or new), and recognition context (whether the target was paired with an old or new host noun). There were two host noun lists (a between-subjects variable) for counterbalancing purposes. Host nouns that were paired with entity nouns on stimuli list 1 were paired with relational nouns on stimuli list 2. The entity nouns and relational nouns were equated on key linguistic features as much as possible. As in Experiment 2, these factors were used to generate a different design for hypothesis testing, as described below. For the subjects analysis, the dependent measures were hits to old targets and false alarms to new targets; these were used to compute d-prime. For the items analysis, the dependent measure was percent correct.

Participants

Two hundred and seventeen undergraduates from the university participant pool participated in exchange for either partial course credit or monetary compensation. (Although we intended to recruit about the same number of participants as in the prior studies, a miscommunication during data collection led to a greater number of participants than planned.)

Materials
All relational nouns and entity nouns were selected from the MRC database (Coltheart, 1981). Nouns that were also used frequently as verbs (e.g., alert) were excluded from the list. Two graduate students in linguistics rated these nouns on the dimension of relationality, based on both the Fetch Test and judgments of entailed arguments (see Supplementary Materials for instructions). We selected nouns that both raters agreed were either entity nouns or relational nouns. The study materials consisted of 118 conceptual combination phrases: 96 experimental items, formed by pairing a ‘host noun’ with an entity or relational noun, and 22 filler items, 10 of which were conventional combinations (e.g., carrot cake). Host nouns were always entity nouns and belonged to one of three categories: animal, plant, or geological formation (e.g., meadow or geyser). Each host noun was used only once on the encoding list for a given participant. We created two stimulus sets in order to control for any effects that a particular host noun might have on encoding and recognition. In other words, if a particular host noun appeared with an entity noun in stimulus set A, it would be paired with a relational noun in stimulus set B.

The recognition materials included four types of conceptual combinations for each type of noun (entity and relational): 12 old conceptual combinations seen during the ratings task ($H_{old}E_{old}$ or $H_{old}R_{old}$), 12 combinations of old host nouns with new entity or relational nouns ($H_{old}E_{new}$ or $H_{old}R_{new}$), 12 combinations of new host nouns with old entity or relational nouns ($H_{new}E_{old}$ or $H_{new}R_{old}$), and 12 completely new combinations ($H_{new}E_{new}$ or $H_{new}R_{new}$). Thus, there were 96 total phrases at test: 48 for entity nouns and 48 for relational nouns.

In constructing the materials, we aimed to equate features that may affect recognition across entity and relational nouns. Using the MRC database (Coltheart, 1981) and frequency norms from Brysbaert and New (2009), we constructed lists of entity and relational nouns such that the distributions were as equivalent as possible with respect to concreteness ($M_{Entity} =$
429.06, $SD = 103.45$, $M_{Relation} = 395.33$, $SD = 88.57$; $t(94) = 1.72$, $p = 0.09$; log10CD

correlated with behavioral data), ($M_{Entity} = 2.611$, $SD = 0.37$, $M_{Relation} = 2.695$, $SD = 0.41$; $t(94) = 1.05$, $p = 0.30$; word length ($M_{Entity} = 7.21$, $SD = 2.51$, $M_{Relation} = 7.42$, $SD = 1.98$; $t(94) = 0.45$, $p = 0.66$), and familiarity ($M_{Entity} = 523.69$, $SD = 46.78$, $M_{Relation} = 518.96$, $SD = 43.63$; $t(94) = 0.51$, $p = 0.61$) (by independent samples t-tests). These characteristics were also used as covariates in the items analysis.

**Procedure**

Participants completed the computer-based ratings task for the 118 conceptual combinations. As before, participants were asked to encode the phrases naturally and to rate the difficulty of constructing an interpretation on a 1-7 scale (see Supplementary Materials for instructions). After a 20-minute unrelated filler task, they received a surprise recognition test. They saw 96 conceptual combinations (48 with entity nouns and 48 with relational nouns) with one word in each pair indicated as the target. Their task was to judge whether the target word had been seen in the original ratings task.

**Results and Discussion**

The results of the recognition test were analyzed using a 2 (host noun list, between-subjects) x 2 (noun focus: entity or relational, within-subjects) x 2 (context: old or new, within-subjects) design. For each participant, we computed measures of recognition accuracy for entity and relational nouns in both new and old contexts. For targets in an old context (i.e., paired with an old host noun), we computed the proportion of hits for old target nouns and the proportion of false alarms to new target nouns (see Table 5). Likewise, to measure recognition accuracy for targets in a new context (paired with a new host noun), we computed the proportion of hits to old
RELATIONAL CATEGORIES ARE MORE MUTABLE

target nouns and the proportion of false alarms to new target nouns (also in Table 5). These results were used to compute participants’ d-prime for old and new contexts.

[Please place Table 5 about here]

There was no main effect of host noun pairing list, $F(1, 215) = .26, p = .61$, so we collapsed the two stimulus lists. A repeated measures ANOVA using noun focus and context as within-subject variables revealed that, as predicted, participants showed greater recognition accuracy for entity nouns ($M = 1.06, SD = .75$) than for relational nouns ($M = .79, SD = .73$), $F(1, 216) = 46.75, p < .001, \eta^2_p = .18$ across old and new contexts. Recognition accuracy was also higher for nouns seen in an old context ($M = 1.16, SD = .74$) than in a new context ($M = .69, SD = .70$), $F(1, 216) = 113.56, p < .001, \eta^2_p = .35$. However, our primary interest was whether a change in context would be more adverse for relational nouns than for entity nouns. Indeed, there was a reliable interaction of noun type and context, $F(1, 216) = 4.63, p = .03, \eta^2_p = .02$.

Recognition accuracy for relational nouns in a new context ($M = .52, SD = .63$) was substantially lower than in an old context ($M = 1.07, SD = .73$), $F(1, 216) = 92.41, p < .001, \eta^2_p = .30$. The difference in recognition accuracy between entity nouns in new context ($M = 0.87, SD = .72$) and an old context ($M = 1.25, SD = .73$) was not as great, $F(1, 216) = 42.31, p < .001, \eta^2_p = .16$ (see Figure 2). These results are very similar to those of Experiment 2; however, in Experiment 3 the increased power from greater sample size allows the interaction to emerge.

[Please place Figure 2 about here]
To evaluate whether this effect held across items, we ran an ANCOVA across items. The dependent variable was proportion correct responses (endorsing old nouns as old and rejecting new nouns as new), and the independent variables were noun focus (entity or relational), recognition context (old host or new host) and target status (whether the tested noun was old or new). After controlling for the effects of log10CD frequency, concreteness, familiarity and word frequency, the analysis revealed that entity nouns were more likely to be identified correctly at test ($M = .68, SD = .10$) than relational nouns ($M = .63, SD = .10$), $F^2(1, 84) = 3.90, p = .05, \eta^2_p = .04$. In addition, nouns appearing in an old context at recognition were identified more successfully ($M = .69, SD = .10$) than nouns in a new context ($M = .61, SD = .10$), $F^2(1,84) = 16.51, p < .001, \eta^2_p = .16$. We also found that new nouns were more likely to be identified correctly ($M = .72, SD = .10$) than old nouns ($M = .59, SD = .10$), $F^2(1, 84) = 46.87, p < .001, \eta^2_p = .36$. However, we did not find the predicted noun focus x recognition context interaction, $F^2(1,84) = .28, p = .60, \eta^2_p = .003$. Word length was the only covariate with a significant relationship to percent correct, $F^2(1,84) = 6.03, p = .02, \eta^2_p = .07$. We are not sure why the interaction failed to appear in the items analysis, given that relational nouns were less well remembered than entity nouns and the interaction appears in the analysis over participants. It could be that some uncontrolled characteristics of our stimuli diluted this effect.

The results from the third experiment generally support the conclusions from the earlier studies. Even when abstractness and other relevant stimuli attributes are controlled, memory for relational nouns is poorer than that for entity nouns, across both participant and items analyses. Further, consistent with the idea that relational nouns are more mutable than entity nouns, analysis across participants shows that relational nouns are more dependent upon their encoding
context than are entity nouns: recognition accuracy for relational nouns suffers more in a novel context than does recognition of entity nouns.

**General Discussion**

Taken together, these experiments support two related claims. The first is that relational nouns are more mutable than entity nouns during encoding—that is, that the interpretation of relational nouns is more influenced by context than is that of entity nouns. The second, related claim is that recognition accuracy for relational nouns is more disadvantaged in a new context than is that for entity nouns. Further, the greater mutability of relational nouns does not result simply from their greater abstractness. Rather, it stems at least in part from the core property of conveying relational meaning.

Our central hypothesis is that when people interpret an ER phrase, they tend to adapt the relational noun to fit the entity noun. Consistent with this hypothesis, in a whole-phrase recognition test (Experiment 1), people were more likely to misjudge a phrase to be “old” if the entity noun was old, (regardless of the relational noun). Relatedly, participants produced more false alarms to test phrases comprising old entity nouns and new relational nouns than to test phrases with new entity nouns and old relational nouns. That is, people were able to recognize a change in entity noun regardless of its relational noun context. However, their recognition accuracy of relational nouns was influenced by its entity noun context.

The results of Experiment 2 also bear out the mutability hypothesis. Participants showed greater ability to discriminate new from old entity nouns than they did for relational nouns. In addition, relational nouns appeared to be more disadvantaged in a novel context than were entity nouns, although this noun x context interaction was not statistically significant. These results suggest that entity nouns were given a stable encoding that could later be recognized in a
RELATIONAL CATEGORIES ARE MORE MUTABLE

relatively context-independent way. In contrast, the encoding of relational nouns was less stable and more influenced by the entity noun with which they occurred. Finally, the findings of Experiment 3 support and extend those of the first two studies. By equating the concreteness of the entity and relational nouns of interest—as well as controlling for many other features that may influence linguistic processing—and by equating the host nouns in the conceptual combinations, we were able to isolate the effect that relationality alone has on mutability. Of course, it is plausible that in natural discourse, relational nouns tend to be more abstract than entity nouns, and that this also contributes to their mutability. But, as discussed below, we suggest that the relation between relational mutability and abstractness is inherent in the nature of language processing.

Comparing relational nouns with verbs. The pattern of findings for relational nouns versus entity nouns parallels that for verbs versus nouns reported by Kersten and Earles (2004). They found better recognition accuracy for nouns than for verbs and a greater drop in accuracy for verbs than for nouns under changes in context. Even when participants were instructed during the encoding phase to try to remember the verbs (Experiment 3), recognition memory was poorer for verbs, and more damaged by changes in context. They noted that their findings could follow from Gentner’s (1981; Gentner & France, 1988) hypothesis that verbs are more contextually mutable than nouns. We propose that an analogous conclusion holds for relational nouns as compared to entity nouns.

As noted earlier, relational nouns share many characteristics with verbs (Gentner, 2005; Gentner & Kurtz, 2005). Relational nouns, like verbs, are centered around patterns of relations, including external relations; and, like verbs, they take arguments (Asmuth & Gentner, 2005). Further, just as verbs are acquired later than nouns, relational nouns are acquired later than entity
nouns (Clark, 1973; Gentner, 2005; Gentner et al., 2011; Gentner & Rattermann, 1991; Hall & Waxman, 1993).

The link between relational nouns and verbs is consistent with accounts of verb meaning that stress the importance of relationality in distinguishing nouns and verbs. For example, in Langacker’s (1987, 2008) account, nouns name things (which are non-relational), while verbs name relationships. Likewise, Croft (1991, 2001) argues that nouns and verbs are linguistically universal categories that differ in both semantic class and pragmatic function. Nouns are the pairing of the semantic class of objects (which are non-relational) and the pragmatic act of reference, while verbs are the pairing of the semantic class of actions (which are relational) and the pragmatic act of predication (i.e., asserting something about another constituent).

Further evidence that relational nouns may be akin to verbs comes from a comprehensive review of behavioral and neuropsychological studies of nouns and verbs (Vigliocco et al., 2011; see also Vinson & Vigliocco, 2002). Drawing on the typological generalizations of Croft and others, the authors noted that the categories of noun and verb are not merely grammatical but also semantic/pragmatic—nouns tend to refer to things, and verbs to actions or events. But they also noted that this correlation is not perfect—for example, that nouns can refer to events or actions. (We would also emphasize that nouns can also refer to relational patterns, as in the example of robbery discussed earlier.)

In their review, Vigliocco et al. (2011) distinguished these semantic factors from the purely grammatical contrast between nouns and verbs. Their results are illuminating. For example, they find that across many studies, low-imageable nouns, and nouns referring to actions/events, often pattern with verbs rather than with other nouns. As Vigliocco et al. (2011; p. 411) note, “Imageability . . . correlates with the distinction between objects and actions/events,
as objects typically tend to be more imageable than events.” In the present research, we find that relational nouns are very likely to be low-imageable nouns. We suggest that at least some of the nouns classified as action nouns in Vigliocco et al.’s studies may be relational nouns, and that this may contribute to their patterning like verbs. For example, in the categorized lists of stimuli given by Vinson and Vigliocco (2002), the list of action nouns includes many that we would classify as relational nouns (e.g., repair, donation, threat, arrival), along with others that refer simply to actions (e.g., sneeze, tremble, scream, sparkle). Although Vigliocco and colleagues do not invoke relationality, their pattern of results is consistent with the idea that relational nouns pattern like verbs. Thus we would endorse their proposal that the grammatical classes of noun and verb have a semantic foundation, and further propose that one important aspect of the semantic difference is relationality.

For our purposes, a key parallel between relational nouns and verbs is that they tend to act as predicates, conveying information about other elements of the sentence. (Croft, 1991, 2001; Langacker, 1987, 2008). For example, although the verb to give normally refers to a transfer of possession, it can convey a transfer of ideas, as in “She gave him the idea for his thesis.” Likewise, the relational noun barrier in a concrete context may refer to a wall or a fence; but in a context such as “a barrier to success” it can refer to insufficient education, ill health or a bad attitude. One way to think of this is that the job of a verb, or a relational noun, is to connect the concepts that are its arguments; and for this to happen, its meaning may need to shift across context. Because the meaning of a relational noun is strongly dependent on its external relations, it is more likely to experience a shift in meaning in different contexts than an entity noun defined by primarily intrinsic properties (See Gentner & France, 1988, for a more detailed version of this discussion with respect to verbs). Indeed, there is evidence that such a shift in representation also
occurs for action verbs when they are used metaphorically. In many studies, figurative uses of verbs (e.g., *The man fell under her spell*) engage secondary areas anterior to primary motion cortex, in contrast to uses denoting literal action (Chatterjee, 2008). These secondary anterior areas may process more abstracted conceptual representations of motion (Chatterjee, 2008, 2010; Jamrozik et al., 2016).

**Implications for conceptual representation.** Our findings are in agreement with recent theories that propose that there are different kinds of categories (e.g., Cree & McRae, 2003; Medin et al, 1997; Moss & Tyler, 2003; Tyler et al., 2000; Vinson & Vigliocco, 2002.) We would further agree with three tenets of Tyler et al.’s (2000) Conceptual Structure theory. The first point of agreement is that componential representations are needed to capture word meaning. There is considerable evidence, both linguistic and psycholinguistic, for componential representations of verbs (e.g., Fillmore, 1971; Gentner, 1981b; Gentner & France, 1988; Langacker, 1987, 1988; Talmy, 1983), and the extension to nouns—especially to relational nouns—has considerable appeal (See Medin & Shoben, 1988; Murphy & Medin, 1985; Rips, 1989). A second point of agreement, broadly shared among many accounts, is that concepts differ in their profiles of components. The third tenet of Conceptual Structure theory is that correlations among features are important and are more critical for some kinds of concepts than for others. While we agree with this, we would go further. We maintain that for many kinds of categories, not only correlations but *relations* are critical—relations that have specific semantic content that goes beyond co-occurrence. This position seems broadly compatible with the emphasis on internal structure in Conceptual Structure theory. For example, consider the paraphrases produced for the phrase “a box examination”:

“studying the contents of a container”
"a test which analyses the ability to think within certain restrictions"

"when the packaging for a product is tested"

"TSA airline safety checks"

In all of these, participants have preserved the idea that an examination involves the relation of checking some other entity for adequacy. Yet the particular combinations of features participants invoke in applying this relational structure to the entity noun box are quite varied. It is not clear how feature correlations could capture this pattern. We suggest that an adequate account of conceptual structure will require representations that include relational structure.

Relationality, mutability and abstraction. We have shown that the greater mutability of relational nouns does not depend on a difference in abstractness. Nonetheless, the strong association between abstractness and relationality commands attention. We note informally that in constructing the materials for Experiment 3, we found it quite difficult to find concrete relational nouns and abstract entity nouns; in contrast, concrete entity nouns and abstract relational nouns are abundant across frequency levels. (In that sense, the materials used in Experiments 1 and 2 may be more representative of ordinary language.) We suggest that this correlation between relationality and abstractness is not a coincidence: rather, it signals a close connection between relationality and abstraction processes. For example, Zharikov and Gentner (2002) traced the evolution of a set of relational nouns using the Oxford English Dictionary. They found that many relational nouns emerged historically as figurative extensions of concrete physical concepts. Some of these extensions preserve the relational structure associated with the original concrete entity, while abandoning the concrete properties. For example, sanctuary once meant a religious edifice, which had the relational property that people could seek safety within. Over time, it developed an abstract, purely relational sense of a place of safety from peril.
Likewise, *bridge* once meant a particular contraption for spanning a river or stream; after a series of extensions, its uses now include written passages that connect parts of texts, loans that cover gaps in funding, and so on—in short, anything that connects two other things or ideas. We suggest that abstract ideas often arise from the process of incrementally extending existing concepts in such a way that they come to convey only an abstract relational structure. This conjecture is supported by the results of a study of which kinds of words are most likely to be used metaphorically. Jamrozik, Sagi, Goldwater, and Gentner (2013) asked people to rate the metaphoricity of uses of verbs, relational nouns and entity nouns from a corpus of natural speech. Metaphoricity ratings were highest for verbs, followed by relational nouns, with entity nouns lowest.

We propose that many abstract concepts are the result of extending concrete entity concepts to become relational concepts (Bowdle & Gentner, 2005; Jamrozik et al., 2016; Wolff & Gentner, 2011). This suggests some specific routes for language change (Heine, Claudi, & Hünnemeyer, 1991; Traugott & Dasher, 2002). This framing may also shed light on the nature of abstractness. Burgoon, Henderson, and Markman (2013) have documented the many ways in which the idea of abstractness is used in psychology. It is possible that at least part of the abstract-concrete distinction should be seen as a distinction between relational concepts versus entity concepts. This would offer a new perspective on the phenomenology of abstract concepts—for example, that they are later in acquisition and harder to remember. It is possible that some of these differences stem from the difference between relational concepts and entity concepts, and not simply from abstractness *per se*.

**Conclusion.** Understanding the representation and processing of relational concepts is critical for accounts of language and cognition. The finding that relational nouns are more
mutable than entity nouns has implications for sentence processing and for the conceptual structure of the lexicon, as well as for metaphor and language change. By comparing relational nouns with entity nouns we can gain insight into the nature of relational language and abstract thought.
Author notes

This research was supported in part by a grant from the Office of Naval Research, award number N00014-02-10078, to the second author. We thank Mike Dickey and Satoru Suzuki for helpful discussions of the issues and Micah Goldwater for suggestions on a prior draft of this paper. We also thank Katherine Braun for her inestimable help in organizing the studies and the Northwestern Analogy and Similarity Group for many thoughtful discussions.
RELATIONAL CATEGORIES ARE MORE MUTABLE

References


RELATIONAL CATEGORIES ARE MORE MUTABLE


RELATIONAL CATEGORIES ARE MORE MUTABLE


Footnotes

1These analyses were done using word frequencies from Kucera and Francis (1967). We repeated the covariate analyses using log10CD frequency norms from Brysbaert & New (2009) in these analyses and found qualitatively the same results.
Table 1

*Experiment 1: Possible Recognition Test Combinations*

<table>
<thead>
<tr>
<th>Encoding phrase:</th>
<th>Entity-Relational order</th>
<th>Relational-Entity order</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a truck limitation</td>
<td>a limitation truck</td>
</tr>
<tr>
<td>Phrase at test:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_{\text{old}}R_{\text{old}}/R_{\text{old}}E_{\text{old}}$</td>
<td>a truck limitation</td>
<td>a limitation truck</td>
</tr>
<tr>
<td>$E_{\text{old}}R_{\text{new}}/R_{\text{new}}E_{\text{old}}$</td>
<td>a truck threat</td>
<td>a threat truck</td>
</tr>
<tr>
<td>$E_{\text{new}}R_{\text{old}}/R_{\text{old}}E_{\text{new}}$</td>
<td>a book limitation</td>
<td>a limitation book</td>
</tr>
<tr>
<td>$E_{\text{new}}R_{\text{new}}/R_{\text{new}}E_{\text{new}}$</td>
<td>a book threat</td>
<td>a threat book</td>
</tr>
</tbody>
</table>
RELATIONAL CATEGORIES ARE MORE MUTABLE

Table 2

*Experiment 1: False alarms*

<table>
<thead>
<tr>
<th></th>
<th>ER</th>
<th></th>
<th>RE</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>FA $E_{oldR_{new}}$</td>
<td>0.28</td>
<td>0.21</td>
<td>0.33</td>
<td>0.28</td>
<td>0.30</td>
<td>0.24</td>
</tr>
<tr>
<td>FA $E_{newR_{old}}$</td>
<td>0.17</td>
<td>0.18</td>
<td>0.21</td>
<td>0.21</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>FA $E_{newR_{new}}$</td>
<td>0.08</td>
<td>0.15</td>
<td>0.14</td>
<td>0.13</td>
<td>0.10</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Table 3

*Experiment 2: Hits, False Alarms, and d’*

<table>
<thead>
<tr>
<th></th>
<th>Entity</th>
<th>Relational</th>
<th>Across focus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td><strong>Old context</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hits</td>
<td>0.73</td>
<td>0.18</td>
<td>0.74</td>
</tr>
<tr>
<td>FA</td>
<td>0.27</td>
<td>0.15</td>
<td>0.39</td>
</tr>
<tr>
<td>d'</td>
<td>1.37</td>
<td>0.66</td>
<td>1.05</td>
</tr>
<tr>
<td><strong>New context</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hits</td>
<td>0.59</td>
<td>0.22</td>
<td>0.50</td>
</tr>
<tr>
<td>FA</td>
<td>0.17</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td>d'</td>
<td>1.32</td>
<td>0.84</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>Across context</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d'</td>
<td>1.34</td>
<td>0.75</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Table 4

*Experiment 3: Possible Recognition Test Combinations*

<table>
<thead>
<tr>
<th>Host/Target</th>
<th>Host-Relational</th>
<th>Host-Entity</th>
<th>‘Yes’ response</th>
</tr>
</thead>
<tbody>
<tr>
<td>old/old</td>
<td>marigold sister</td>
<td>carnation substance</td>
<td>Hit</td>
</tr>
<tr>
<td>old/new</td>
<td>marigold victim</td>
<td>carnation article</td>
<td>FA</td>
</tr>
<tr>
<td>new/old</td>
<td>albatross sister</td>
<td>puddle substance</td>
<td>Hit</td>
</tr>
<tr>
<td>new/new</td>
<td>albatross victim</td>
<td>puddle article</td>
<td>FA</td>
</tr>
</tbody>
</table>
Table 5

Experiment 3: Hits, False Alarms, and \( d' \)

<table>
<thead>
<tr>
<th>Entity</th>
<th>Relational</th>
<th>Across focus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Old context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hits</td>
<td>0.75</td>
<td>0.16</td>
</tr>
<tr>
<td>FA</td>
<td>0.33</td>
<td>0.17</td>
</tr>
<tr>
<td>( d' )</td>
<td>1.25</td>
<td>0.73</td>
</tr>
<tr>
<td>New context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hits</td>
<td>0.52</td>
<td>0.21</td>
</tr>
<tr>
<td>FA</td>
<td>0.22</td>
<td>0.14</td>
</tr>
<tr>
<td>( d' )</td>
<td>0.87</td>
<td>0.72</td>
</tr>
<tr>
<td>Across contexts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( d' )</td>
<td>1.06</td>
<td>0.75</td>
</tr>
</tbody>
</table>
Figure 1. Mean $d'$ of entity and relational nouns in old and new contexts in Experiment 2. Error bars depict standard error.

Figure 2. Mean $d'$ of entity and relational nouns in old and new contexts in Experiment 3. Error bars depict standard error.