Learning Relational Concepts through Unitary versus Compositional Representations

Daniel Corral (daniel.corral@colorado.edu) & Matt Jones (mcj @colorado.edu)
Department of Psychology and Neuroscience, University Colorado Boulder,
Boulder, CO 80309 USA

Abstract

Current theories of relational learning on structure mapping emphasize the importance of compositional representations, based on the concept's components and the relations among them. We consider the possibility that relational concepts can also be represented *unitarily*, whereby the concept is a property of the stimulus as a whole. The distinction between compositional and unitary representations of relational concepts is a natural consequence of structure-mapping theory, but its psychological implications have not been explored. We report two experiments in which we examine how encouraging subjects to represent relational concepts compositionally versus unitarily affects learning on classification- and inference-based category learning tasks. Our findings show that unitary representations lead to better learning than compositional representations, especially for the inference task. We conclude that unitary representations incur less cognitive load than structural alignment of compositional representations, and thus may be the default for everyday relational reasoning.

Keywords: Relational Learning; Relational Structure; Concept Representation; Category Learning; Inference.

Introduction

On a daily basis, people encounter many complex concepts that are defined by a *relational structure* – the specific pattern in which two or more objects are bound together by interconnected relations (Corral & Jones, 2014). For instance, consider a simple scenario in which a dog chases a cat. In this example, the dog and the cat share a specific relationship with one another, such that it is the dog that fills the role of the chaser and the cat fills the role of being chased (*chase*(*dog*, *cat*)). Critically, this structure is different from a scenario in which a cat chases a dog (*chase*(*cat*, *dog*)). These types of concepts differ from those that are defined by features, which can be identified by the presence of a given set of attributes (Estes, 1986). For example, a bird might be identified by the presence of certain prototypical features, such as {*feathers*, *beak*, *wings* …}. Although feature-based representations can provide extensive knowledge about a given scenario, they do not convey structural information (Markman, 1999). Thus, a feature-based representation does not allow one to readily distinguish a simple scenario in which a dog chases a cat from an instance in which a cat chases a dog (Markman & Gentner, 2000), as both would be represented as an unstructured set: {*dog*, *cat*, *chase*}.

The ability to recognize and reason about structured concepts has been posited to be one of the cornerstones of human cognition (Penn, Holyoak, & Povinelli, 2008). According to structure-mapping theory, the dominant theory of relational learning, structured concepts are acquired via structure mapping, wherein the elements of two analogous scenarios are put into alignment in a way that preserves their common roles. For example, in the hypothetical scenarios described below, the dog in the first scenario maps to the cat in the second scenario because both fill the role of the chaser. Alignment of two scenarios highlights their common structure and facilitates abstraction of new relational concepts (Gentner, 1983; Hummel & Holyoak, 2003).

Importantly, structure-mapping theory[1] makes the implicit assumption that a relational concept can be represented in two fundamentally different ways: (1) as a system of relations, with meaning derived both from the identities of those relations and from how they are interconnected by shared role-fillers (Corral & Jones, 2014); or (2) as a primitive, atomic relation that is explicitly represented. We refer to these as *compositional* and *unitary* representations. Although this logical distinction has been noted (Gentner, 1983), its potential psychological implications have largely been neglected.

To elaborate further, the first of these representational assumptions is premised on the idea that representations are constructed from two basic types of building blocks: objects and relations. The second assumption is based on the idea that a relation operates on a set of n objects, that is, for every ordered set of n objects, the relation returns a truth-value indicating whether the objects satisfy the relation. Equivalently, for every ordered set of n objects ($o_1\ldots,o_n$) for which the relation holds, there is an explicit token of that relation: $R(o_1\ldots,o_n)$. We refer to any relation of this sort as a unitary relation.

In recent work, Corral, Kurtz, and Jones (under revision) raise the possibility that subjects might indeed represent some relational concepts unitarily, such that the concept is a component or a property of the stimulus as a whole. This type of representation would lack explicit structure and could be recognized directly in a stimulus, similarly to a feature. This idea is perhaps best exemplified in language comprehension, where people appear to seamlessly understand a multitude of rich relational concepts, without explicitly representing their substructure. For example, consider the concept of *investigation*. An investigation consists of an agent, a given question, the approach the agent takes to answering that question, and the specific

[1] It is important to note that there are numerous domains within cognitive science that formalize representation in various ways. In the present paper, we work within the framework of structure-mapping theory.

pattern of interconnections among these components. Nevertheless, people can likely recognize this concept without explicitly representing its structure. Likewise, a *t*-test involves a complex structure of mathematical elements and relations (as many hapless introductory statistics students will attest), but for experienced scientists it is easily conceived of as a unitary event—one can hear the sentence "I ran a t-test" and immediately comprehend its meaning without needing to invoke the concept's substructure.

The literature on structure-mapping theory has focused on compositional representations, through its emphasis on the alignment process. Furthermore, it has been proposed that people must use compositional representations in order to learn relational concepts (Markman & Gentner, 2000). Compositional representations are computationally expensive (Forbus, Gentner, & Law, 1995) and can place a high strain on working memory (Kintsch & Bowles, 2002). They are also unnecessary for learning feature-based concepts (Markman, 1999), which can be recognized (without regard to structure) by attending to a stimulus' defining attributes (e.g., Nosofsky, 1986). Similarly, relational concepts that are represented unitarily can be explicitly recognized as a global attribute of the given scenario, and thus can be learned in an unstructured manner. Such representations allow for computationally efficient processing (Forbus et al., 1995), and based on principles of cognitive economy, it follows that people should avoid compositional representations and structural alignment whenever a unitary representation and setwise (feature-style) comparisons are adequate.

Evidence from related literatures suggests that people in fact do not use compositional representations as much as might be expected based on structure-mapping theory. One prediction that follows from compositional representations is that people should be able to report the structural elements of the relational concepts they are familiar with. However, despite subjects reporting high confidence in their comprehension of various types of common relational systems (e.g., how helicopters fly), they are often mostly unaware of their structural elements (Keil, 2003; Rozenblit & Keil, 2002). Another prediction from compositional representations is that, because relational structure must be explicitly represented (Kintsch & Bowles, 2002), it should take longer to comprehend and recognize structured information than information that is not structured. However, various studies have found no differences in the time it takes subjects to comprehend structured (metaphors) and non-structured statements (e.g., "*the ball is blue*") (Glucksberg, Gildea, Bookin, 1982). Related work has shown that subjects can often understand metaphors automatically, with minimal explicit processing (Glucksberg, 2003). Taken together, these findings suggest that many relational concepts may not typically be represented compositionally.

Due to the representational flexibility that humans possess (Chalmers, French, Hofstadter, 1992), it seems plausible that relational concepts can be represented both unitarily and compositionally. For instance, a person might represent a concept such as *investigation* based on a global attribute (e.g., an inspection), but can also likely represent its relational substructure when necessary (explicitly representing the agent, question, line of inquiry, and their interrelations). This idea leads to the question of which type of representation people use by default when learning a relational concept. The main hypothesis of the present paper is that, because unitary representations should allow for more efficient processing, subjects will use such representations when they are available. We test this prediction by giving subjects relational category learning tasks and encouraging them to represent the stimuli either compositionally or unitarily. If people typically learn relational concepts from structural alignment, then encouraging subjects to use compositional representations should aid learning. However, if people instead learn more efficiently with unitary representations, than the opposite outcome should be expected.

Half the subjects in our experiments were given a classification task, in which they were shown a series of stimuli and asked to make categorization judgments. Unitary representations seem especially well-suited for such a task, because they should enable subjects to directly recognize the diagnostic property in a stimulus, just as with feature-based categories. The other subjects were given an inference task, in which they were asked on each trial to determine a missing property of a stimulus that was presented together with its category label. Research with feature-based categories has shown that classification and inference learning tend to yield different category representations, with inference tasks encouraging learning of internal category structure, such as correlations among features (Markman & Ross, 2003; Yamauchi & Markman, 2000). This finding suggests that compositional representations should be particularly well-suited for inference learning with relational categories, as such representations highlight the internal structure of stimuli. The inference conditions of our experiments thus provide a more stringent test of our hypothesis that people can learn relational concepts better through unitary representations.

Experiment 1

Experiment 1 examines how providing unitary and compositional descriptions of relational concepts affects learning on classification and inference tasks (description and task type both manipulated between subjects). Subjects were provided either a unitary or compositional hint at the start of learning and again after every third error, in order to assess whether each type of hint can improve learning. Control groups who were given no hints were also included in order to assess baseline performance in both tasks.

The stimuli used in this study were taken from Corral, et al. (under revision), which were adopted and modeled after those used by Rehder and Ross (2001). A stimulus consisted of three sentences, each of which describes a different component of a machine that works to remove waste material: (1) the location of where the machine operates, (2) the waste material the machine removes, and (3) the instrument the machine uses.

Stimuli were sampled from two categories: coherent and incoherent. Each category consisted of 18 exemplars. The categories were determined by how a machine's components were related to one another. For exemplars from the coherent category, the machine's instrument is suited for collecting the waste material that the machine works to remove, which can be found in the location where the machine operates. Consider the following example: "Operates on the seafloor, works to remove lost fishing nets, and has a hook." This exemplar is coherent because of the secondary relations among the machine's component parts (presumed to be known by subjects), such that lost fishing nets can be found on the seafloor and a hook can be used to retrieve lost fishing nets. In contrast, exemplars from the incoherent category do not satisfy either of these second-order relations (i.e., the machine's tool cannot be used to collect the machine's target waste material and that material cannot be found where the machine operates). Non-Morkels were thus made to be as incoherent as possible so as to maximally differentiate the categories and better facilitate learning of the task. Figure 1 illustrates the abstract relational structure of the two categories.

Half of the subjects completed an A/¬A classification task (in which each stimulus was to be categorized as either a category member or a nonmember), and the other half completed an inference task. On each trial, the subject was presented a single stimulus and asked to make an inference or classification judgment (depending on the condition). After the response, the subject was shown whether the response was correct along with the correct answer.

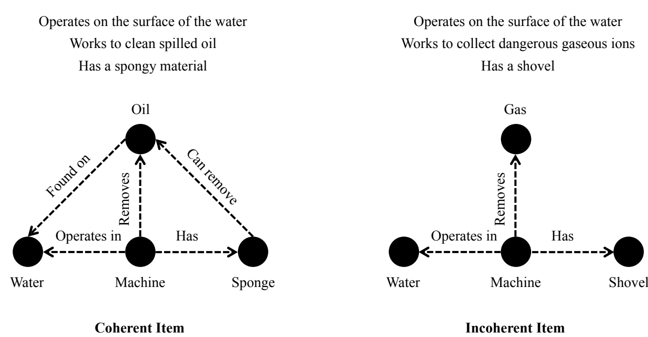

Figure 1. Illustration of the relational structure for items in the coherent and incoherent categories in Experiment 1. The structures differ in that coherent items satisfy the relations indicated by diagonal lines: the machine's implement can remove the target, and the target is found in the machine's location. Recreated from Corral et al. (under revision).

Method

Two hundred eighteen undergraduates from the University of Colorado Boulder participated for course credit in an introductory psychology course. Subjects were randomly assigned to six conditions. Type of hint (compositional vs. unitary vs. control) was crossed with task type (classification vs. inference).

Subjects were told that they would be shown short descriptions of various types of cleaning machines, some of which were made by the Morkel Company (coherent category) and some were not (incoherent category). Subjects were provided a positive example of a Morkel (randomly selected) and told that all Morkels share a certain commonality and it was their job to figure out what it was.

Subjects in the unitary condition were shown the following hint: "On each trial try to think about how "well suited" the machine is for performing its task. Keep in mind that consumers say machines from Morkels are built "intuitively" in a way that makes sense." This hint was intended to shift subjects' attention toward finding a global attribute of the stimulus and away from the explicit relationships among its components. Using this hint, it is possible for subjects to learn how to distinguish the categories without explicit knowledge of their relational structure. This hint can therefore be said to encourage subjects to represent each stimulus unitarily.

Subjects in the compositional condition where shown the following hint: "On each trial try to think about the specific manner in which the machine's 1^{st} property relates to its 2^{nd} and 3^{rd} properties, as well as how its 2^{nd} property relates to its 3^{rd} property." This hint was intended to focus subjects' attention on the relationships among the component parts of the stimulus, and thus to encourage them to represent the stimulus compositionally.

Subjects were presented the appropriate hint during the initial task instructions, after the first trial, during rest breaks, and following every third error the subject committed (on a blank screen after corrective feedback was shown). Subjects were asked to read the hint carefully and press the spacebar when they were ready to continue. Subjects in the control group were not shown a hint and were instead asked to continue to try their best; this reminder was presented on every third error the subject committed and on rest breaks.

Each subject completed 72 trials. The order in which the items were presented was randomized for all subjects. In each block of 18 trials, all 18 stimuli appeared in a random order. After each block, subjects were given a self-paced rest break and were shown the proportion of correct responses they answered correctly over those trials, along with the number of trials they had completed and the number that remained.

On each trial in the classification condition, a single, complete stimulus was presented and the subject was asked to type "A" if the machine was a Morkel or "L" if it was not. On each trial in the inference condition, the category label for a stimulus was shown (Morkel or non-Morkel) directly above an incomplete stimulus consisting of two of its three components (i.e., sentences). Below the stimulus were two response options, one of which was the missing component and the other was a lure. The component the subject was asked to infer (i.e., implement, target material, or location) was randomly selected on each trial. Subjects were asked to select which was the missing component by typing "A" if the correct choice was the top option or "L" if it was the bottom option. The order in which the two options were presented was randomized on every trial. For items that were Morkels, the correct response was the option that shared secondary relations with the given stimulus components. The lure did not share secondary relations with either of the stimulus components. For items that were non-Morkels, the correct response was the component that did not share any secondary relations with either of the stimulus components. The accompanying lure shared at least one secondary relation with one of the stimulus components. Figure 2 shows an example trial from the inference condition.

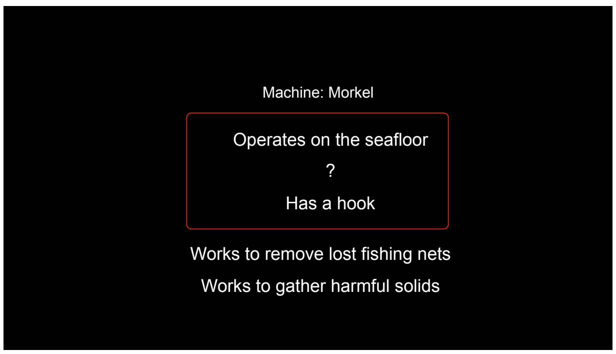

Figure 2. Example of a stimulus display from the coherent category (Morkels) from the inference task in Experiment 1.

Results & Discussion

Figure 3 shows average learning curves for subjects in each group. An ANOVA was conducted to examine differences in performance among groups. The analysis showed a main effect of hint, $F(2, 212) = 42.14$, $p < .0001$, $MSE = .014$, and an interaction, $F(1, 212) = 8.90$, $p = .0002$, $MSE = .014$, indicating that the main effect of hint depends on the type of task that subjects completed. On the classification task, control subjects ($M = .61$, $SE = .017$) were outperformed by subjects in the compositional ($M = .775$, $SE = .014$; $p < .0001$) and unitary groups ($M = .83$, $SE = .016$; $p < .0001$). In the inference condition, only subjects who received a unitary hint ($M = .716$, $SE = .012$) performed better than control subjects ($M = .585$, $SE = .012$; $p < .0001$), as no differences were observed between subjects who were presented a compositional hint ($M = .587$, $SE = .011$) and subjects in the control group.

Planned *t*-tests were conducted to compare the unitary and compositional groups, separately for each task. On the classification task, subjects in the unitary condition ($M = .83$, $SE = .014$) outperformed subjects in the compositional condition ($M = .775$, $SE = .014$), $t(71) = 1.85$, $p = .068$, $d = .45$. This same pattern was observed in the inference condition (unitary $M = .716$, $SE = .012$; compositional $M = .587$, $SE = .012$), $t(67) = 5.28$, $p < .0001$, $d = 1.29$. An additional 2 (unitary vs. compositional) × 2 (classification vs. inference) ANOVA was conducted, which excluded control subjects. This analysis revealed an interaction, $F(1, 138) = 4.01$, $p = .047$, $MSE = .013$, indicating that the unitary advantage was stronger in the inference task than in the classification task.

Taken together, the findings presented here suggest that unitary and compositional representations can both be used to acquire relational concepts. However, subjects who were encouraged to represent the stimuli unitarily showed more robust learning than subjects who were encouraged to represent the stimuli compositionally, especially in the inference task. These findings thus provide support for our main hypothesis that, when both types of representations are available, subjects learn better with unitary than with compositional representations.

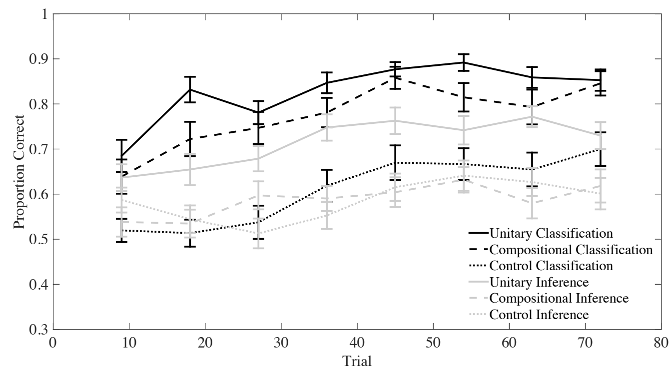

Figure 3. Average learning curves and standard errors across blocks of nine trials for each condition in Experiment 1.

Experiment 2

Experiment 2 builds on the findings from Experiment 1 and examines how category learning is affected when subjects represent a relational concept one way (either unitarily or compositionally) and are subsequently made aware of an alternative representation. Experiment 2 used the stimuli from Experiment 1, and all subjects performed the classification task. All subjects were either provided a unitary or compositional hint prior to the start of learning. For half of the subjects, the hint was changed after the 18th trial (i.e., the unitary hint was replaced with the compositional one and vice versa). For the other half of subjects, the hint they were shown remained the same throughout the study. These latter conditions were identical to the unitary and compositional classification conditions in Experiment 1.

Method

One hundred fifty-seven subjects were randomly assigned to four conditions: unitary/switch ($N = 40$), compositional/switch ($N = 39$), unitary/no-switch ($N = 39$), and compositional/no-switch ($N = 39$). After the 18th trial (i.e., in the first rest break), the screen was cleared and subjects in the switch conditions were shown a prompt that notified them that Morkels could be represented differently from the initial hint and were shown the other hint. Following the 19th trial, this hint was presented once more and subjects were reminded to use it to try to figure out what constitutes a Morkel. Subjects in the switch conditions were shown this hint for the remainder of the study (i.e., on rest breaks and following every 3rd error), whereas no-switch subjects continued to see the hint they had seen at the beginning. The rest of the procedure was identical to that of Experiment 1.

Results & Discussion

Figure 4 shows average learning curves for subjects in each condition. A *t*-test showed that subjects in the unitary/no-switch condition (M = .79, SE = .017) outperformed subjects in the compositional/no-switch condition (M = .716, SE = .017), $t(76)$ = 2.23, p = .03, d = .50. This finding directly replicates the results from the classification condition in Experiment 1, which showed a unitary learning advantage.

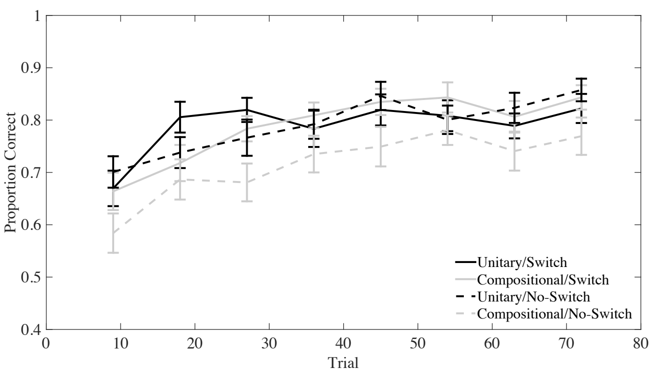

Figure 3. Average learning curves and standard errors across blocks of nine trials for each condition in Experiment 2.

In addition to this analysis, a series of planned comparisons were conducted to examine differences among groups from the point at which subjects were introduced to the other hint (trials 19-72). The first analysis showed that subjects in the compositional/switch condition (M = .82, SE = .017) outperformed subjects in the compositional/no-switch condition (M = .743, SE = .017), $t(76)$ = 2.26, p = .027, d = .51. Additionally, subjects in the unitary/switch condition (M = .802, SE = .018) marginally outperformed subjects in the compositional/no-switch condition, $t(77)$ = 1.77, p = .08, d = .45. However, no differences in performance were observed among any of the three groups that were presented a unitary hint at some point in the study. Thus, it seems that as long as a unitary hint is presented, regardless of whether it is the only hint that is shown or if it is presented before or after a compositional hint, subjects are able to benefit from it. Taken together, these findings support the conclusion from Experiment 1 and suggest that subjects indeed learn better when they rely on unitary representations.

General Discussion

We report two experiments that test how encouraging subjects to represent relational stimuli unitarily or compositionally affects concept learning. The findings from Experiment 1 showed that both types of hints can aid learning on a classification task, but only the unitary hint was a useful learning aid on the inference task. These findings provide support for the idea that subjects can indeed use both types of representations to understand and learn relational concepts, but that unitary representations are as or more effective than compositional ones. This latter conclusion challenges the emphasis on compositional representations at the core of most research on analogical reasoning.

Experiment 2 used only a classification task and was able to replicate the findings from the classification condition in Experiment 1, as subjects who received only a unitary hint outperformed subjects who received only a compositional hint. Furthermore, the results from this study showed that subjects who received a unitary hint at any point in the study (with a compositional hint coming before, after, or not at all) outperformed subjects who did not receive a unitary hint at all. No differences in performance were found among subjects in the groups who received a unitary hint. These results lend more support to the dominance of unitary representations, in that subjects will abandon or ignore suggestions for compositional representations if they have discovered a unitary one.

One surprising finding from Experiment 1 was that the unitary advantage was stronger for the inference task than for classification. The effect size for the inference task was actually quite dramatic (Cohen's d of 1.29). We had predicted that, if anything, the interaction would go in the opposite direction, given that inference tasks encourage learning the relationships among a concept's components (Markman & Ross, 2003; Yamauchi & Markman, 2000). One speculative possibility is that inference learning encourages a top-down approach, in that subjects must reason from the category label to the stimulus, whereas classification encourages a bottom-up approach of reasoning from the stimulus to the category label. Likewise, a unitary representation is top-down in that it embodies a global property of a stimulus that can be used to deduce its internal structure, whereas a compositional representation is bottom-up in that the local structure is explicitly represented and the global property emerges only implicitly from the relational system. Under this view, there might be a congruency effect between the stimulus representation and the processes involved in carrying out the task. In particular, a unitary representation might be more congruent with an inference task, because it facilitates conceiving of a concept by a single attribute that can then be used to infer missing parts of a stimulus.

These speculations aside, the main conclusion of the present studies is that, although relational concepts are defined by the interconnections among their component parts, subjects seem to learn these concepts better when they can be represented unitarily, which might facilitate a global understanding that is easier to discover and use than an explicitly structured one. Furthermore, although compositional-based instruction can help subjects classify a given concept, it might not be optimal for inference-based reasoning.

These findings seem particularly applicable to education and instruction, as they might provide insight into how different types of descriptions for a given relational concept can affect students' representations, as well as how such representations affect learning. Indeed, students are often required to learn various types of structured concepts, and must often engage in both classification and inference. For instance, in mathematics, students must recognize various instantiations of a given problem type, a process that relies on classification, and must also make inferences about how to apply a given solution. These findings thus hold the potential to improve how relational concepts are taught in the classroom.

Furthermore, the present findings have theoretical implications for relational concept learning and representation, and have the potential to affect current theories of analogical reasoning and learning. In particular, research within the theoretical framework of structure mapping (Doumas, Hummel, & Sandhofer, 2008; Hummel & Holyoak, 2003) has placed a heavy emphasis on alignment processes operating on compositional representations, but our findings suggest that subjects more naturally represent such concepts unitarily, and that such representations produce a greater and more robust benefit to learning. During comparison of two scenarios, if the critical information can be represented unitarily, then there is no need for structural alignment, because the two can be recognized through the same sort of processing that is possible with feature-based representations, that is, flat (setwise) comparison to identify which properties they have in common. To be clear, this proposal is not intended to argue against the idea that structural alignment of compositional representations plays a prominent role in the more impressive feats of human reasoning (e.g., creativity or scientific discovery), but rather to point out that in more mundane cases, simpler processes and representations may be involved. Nevertheless, further work is necessary to better understand which conditions facilitate unitary and compositional representations.

Lastly, we note one potential shortcoming of the present studies. Although subjects were encouraged to represent the stimuli unitarily or compositionally, we cannot know for certain whether subjects adopted either of these representations. This issue has historically plagued researchers in this domain of study and highlights the need for improved assessment on concept representation. We welcome suggestions in helping us to address this challenge.

References

Chalmers, D. J., French, R. M., & Hofstadter, D. R. (1992). High-level perception, representation, and analogy: A critique of artificial intelligence methodology. *Journal of Experimental and Theoretical and Artificial Intelligence, 4*, 185-211.

Corral, D., & Jones, M (2014). The effects of relational structure on analogical learning. *Cognition, 132*, 280-300.

Corral, D., Kurtz, K. J. & Jones, M. *Learning relational concepts from within- vs. between-category comparisons*. Under revision.

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*, 1-43.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology, 18,* 500-549.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141-205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155-170.

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences, 7*, 92–96.

Glucksberg, S., Gildea, P., & Bookin, H. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior, 21,* 85-98.

Hummel J. E., & Holyoak K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220-264.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Science*, *7*, 368-373.

Kintsch, W., & Bowles, A. R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, *17*, 249–262.

Markman, A. B. (1999). Knowledge representation*.* Hillsdale, NJ: Erlbaum.

Markman, A. B., & Gentner, D. (2000). Structure-mapping in the comparison process. *American Journal of Psychology*, *113*, 501-538.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129*, 592-615.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

Penn, D. C., Holyoak, K J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31*, 109-178.

Rehder, B. & Ross, B.H. (2001). Abstract coherent concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1261-1275.

Rozenblit, L. R. & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26,* 521-562.

Yamauchi, T., & Markman, A. B. (2000). Inference using categories. Journal of Experimental Psychology: Learning, *Memory and Cognition, 26*, 776-795.