

Comparison Promotes Learning and Transfer of Relational Categories

Kenneth J. Kurtz
Binghamton University

Olga Boukrina
Rutgers University

Dedre Gentner
Northwestern University

We investigated the effect of co-presenting training items during supervised classification learning of novel relational categories. Strong evidence exists that comparison induces a structural alignment process that renders common relational structure more salient. We hypothesized that comparisons between exemplars would facilitate learning and transfer of categories that cohere around a common relational property. The effect of comparison was investigated using learning trials that elicited a separate classification response for each item in presentation pairs that could be drawn from the same or different categories. This methodology ensures consideration of both items and invites comparison through an implicit same–different judgment inherent in making the two responses. In a test phase measuring learning and transfer, the comparison group significantly outperformed a control group receiving an equivalent training session of single-item classification learning. Comparison-based learners also outperformed the control group on a test of far transfer, that is, the ability to accurately classify items from a novel domain that was relationally alike, but surface-dissimilar, to the training materials. Theoretical and applied implications of this comparison advantage are discussed.

Keywords: comparison, categorization, category learning, structural alignment, relational categories

Comparison and categorization are two of the core mechanisms that underlie human learning, understanding, and reasoning. Yet for the most part, they have been studied quite separately (for reviews, see Gentner, Holyoak, & Kokinov, 2001; Gentner & Markman, 1997; Holyoak & Thagard, 1995; Levering & Kurtz, 2010; Murphy, 2002; Ross, Taylor, Middleton, & Nokes, 2008). In the study of categorization, much research has focused on the classification learning paradigm—across a series of trials the learner is visually presented with an item drawn from a training set, then a response is made by choosing which category the item belongs to, and corrective feedback is received. This research approach comes with inherent compromise in terms of ecological validity (one-at-a-time presentation of items that manifest values on a small, fixed set of independent binary attributes for two-choice classification) but has served to reveal a great deal about the processes and representations underlying the learning, organiza-

tion, and use of category knowledge. Patterns of human performance in classification learning have been fit by formal models instantiating psychological constructs including attention-weighted similarity to individual or clustered exemplars, logical rules plus exceptions (see Pothos & Wills, 2011, for broad coverage of these approaches), and compatibility of recoding/decoding schemes in a divergent auto-associative connectionist network (Kurtz, 2007).

But despite the success of this approach, some kinds of categories may be difficult to capture with these kinds of models. A growing body of research centers around relational categories (Gentner & Kurtz, 2005; Markman & Stilwell, 2001)—categories whose membership is determined by common relational structure rather than common intrinsic features. An example is *prize*, which can encompass any reward achieved through an action—be it a toy in a crackerjack box, a blue ribbon, or a ship captured in war. Another example is *conflict*—that is, a state of incompatibility or struggle between opposing forces—which again can be applied to forces ranging from ranchers to wolves to competing ideologies. Relational categories figure prominently in adult discourse, in part because they allow complex relational structures to be “packaged” in a way that permits further predication (Gentner & Kurtz, 2005)—as in “The inevitable *outcome* of the *conflict* between economic policy and social policy is a *stalemate* benefiting neither.” Informal ratings of the 100 highest frequency nouns in the British National Corpus suggest that close to half are relational (Asmuth & Gentner, 2013), and many superordinate categories are relational as well (e.g., *carnivore* and *pet*; Gentner & Asmuth, 2008).

A further reason to study relational categories is that they behave differently from entity categories (those whose members

This article was published Online First February 18, 2013.

Kenneth J. Kurtz, Department of Psychology, Binghamton University; Olga Boukrina, Department of Psychology, Rutgers University; Dedre Gentner, Department of Psychology, Northwestern University.

This research was supported in part by a National Institutes of Health NRSA postdoctoral award and by an Office of Naval Research (# N00014-02-1-0040) award to the third author.

We thank various colleagues and the members of the Learning and Representation in Cognition (LaRC) Laboratory at Binghamton University for helpful contributions to this project.

Correspondence concerning this article should be addressed to Kenneth J. Kurtz, Department of Psychology, Binghamton University, P.O. Box 6000, Binghamton, NY 13902-6000. E-mail: kkurtz@binghamton.edu

share intrinsic properties). Relational categories are slower to be learned by children, who often initially interpret them as entity categories (Gentner, 2005). Relational categories are more likely to be described by ideal features (rather than typical features) than are entity categories (Goldwater, Markman, & Stilwell, 2011), and the most effective standard is an ideal member, not the most typical member (Rein, Goldwater, & Markman, 2010). These differences, coupled with their frequency and semantically interesting behavior, make clear that to arrive at a complete understanding of human categorization, we must look beyond entity categories. Accordingly, relational categories are receiving increased empirical and theoretical attention (e.g., Asmuth & Gentner, 2005; Doumas, Hummel, & Sandhofer, 2008; Gentner, 2005; Gentner, Anggoro, & Klibanoff, 2011; Goldwater & Markman, 2011; Goldwater et al., 2011; Jones & Love, 2007; Kurtz & Gentner, 2001, 2013; Rehder & Ross, 2001; Tomlinson & Love, 2010; Wiemer-Hastings & Xu, 2005).

A key question that arises from these considerations is: How do people learn relational categories? Drawing on research from the study of analogy, the comparison process is a promising candidate mechanism. Comparison-based learning is a process of aligning the relational predicates of two cases so that their common, connected structure is rendered salient and extractable as an abstract knowledge structure (Gentner, 1989, 2010; Gentner & Medina, 1998; Gick & Holyoak, 1983; Hummel & Holyoak, 1997). Such learning can be elicited by juxtaposition of cases in the external environment (Christie & Gentner, 2010; Kotovsky & Gentner, 1996), by applying a label to multiple cases (Namy & Gentner, 2001), or via reminding (e.g., Ross, Perkins, & Tenpenny, 1990). The internal relational structure of a case is assumed to be psychologically encoded as a structured representation rather than a flat feature list or a point in multidimensional space. Comparison highlights common relational structure leading the learner to perceive a more general representational structure—promoting schema abstraction. These mechanisms operate within the structural alignment process articulated in the structure-mapping theory of similarity and analogy (Gentner, 1983; Gentner & Markman, 1997).

The empirical base for comparison-based learning is grounded in evidence of a different kind than is usual in the categorization literature. We know that comparison supports knowledge transfer and improved, more principle-based performance in arenas such as problem solving (Gick & Holyoak, 1983; Kurtz & Loewenstein, 2007; Ross & Kennedy, 1990), learning negotiation strategies (Loewenstein, Thompson, & Gentner, 1999), scientific understanding and insight (Kurtz, Miao, & Gentner, 2001), learning to achieve stability in construction (Gentner, Levine, Dhillon, & Poltermann, 2009), and memory retrieval (Gentner et al., 2009). This range of evidence shows the power of comparison to promote abstraction of common structure from examples that have common relational content bound to distinct surface elements.

This aspect of comparison processing—its capacity to invite the formation of relational schemas—suggests that it may be important for categories that share common relational structure. Goldwater and Markman (2011) found that role categories whose referents play a particular role in a relational schema can be made more salient by comparing two members and by using a common label. Further support comes from developmental evidence concerning children’s word learning. For example, 3- and 4-year-olds

are more likely to extend a new label according to a shared relational pattern than according to an object match if they have compared two examples rather than seeing just one (Christie & Gentner, 2010; see also Gentner et al., 2011; Gentner & Namy, 1999). These findings suggest that the structural alignment process invites a relational encoding that can serve as a foundation for learning.

The goal of this research was to test whether and how comparison-based learning can promote the acquisition of relational categories in adults. We believe that promoting comparison will be especially important in learning relational categories for several reasons: (a) Relational category acquisition depends on abstracting the relational structure shared by members; (b) the relational properties underlying membership are likely not in the default encodings of the stimuli; and (c) productive spontaneous reminders are unlikely since these tend to be driven by surface similarity. A further motivation for this line of inquiry comes from the potential application to pedagogy. A successful comparison-based learning technique would facilitate formal instruction of relationally based, abstract concepts (e.g., *equality*, *density*) in science and math (see Rittle-Johnson & Star, 2009).

Experimental Approach

We began with a classification learning task and investigated whether side-by-side comparison of training items would promote acquisition of novel relational categories. The stimuli were moderately naturalistic line-drawn images depicting “rock arrangements” made by imaginary cultures (see Figure 1 and the Appendix). The exemplars each instantiate a specific, systematic spatial relation and lack any clear reduction to a compositional set of underlying dimension values. A three-category domain was used to avoid biases and artifice inherent in two-choice classification (i.e., perfect accuracy can be based on knowledge of only one category; task demands encourage hypothesis testing for diagnostic features or a decision boundary, as opposed to the acquisition of positively defined concepts). Before continuing to the main experiment, we begin by reviewing results and challenges encountered in preliminary work.

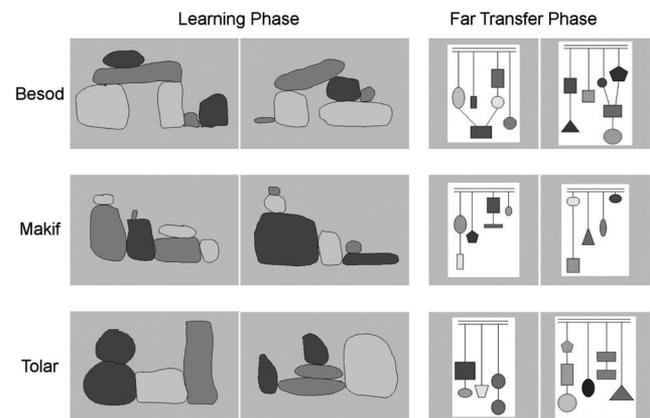


Figure 1. Sample learning and testing materials.

Preliminary Study

In an initial approach to this line of inquiry, Kurtz and Gentner (1998) found that participants learned relational categories via classification more quickly when presented with within-category pairs (jointly classified with a single response) than when presented with the same number of single-item trials. However, this comparison advantage may have resulted from comparison learners receiving twice as many item exposures. To address this concern, we conducted follow-up research using improved materials and controlling for the amount of exposure by cutting in half the number of trials (guess-and-correct cycles) for comparison learners. In the single-item control condition ($n = 46$), learners received 48 training trials—two passes through the training set of three relational categories each consisting of eight rock arrangements (described in detail in the Materials section below and depicted in Figure 1). In the comparison condition ($n = 49$), items were presented in within-category pairs for 24 training trials. This resulted in a total of two presentations per item in both conditions. At test, both groups were asked to classify new and old items in a single-item presentation format without feedback.

The results made clear that the experimental materials were learnable yet sufficiently challenging to avoid ceiling effects (i.e., the underlying relations were nonobvious and free of trivial cues). However, there was no significant difference between groups and therefore no support for the predicted comparison advantage. An initial interpretation might be that with proper controls for item exposure, there is no advantage associated with comparison. However, while the conditions were equated for number of item exposures, this resulted in the single-item control group having the advantage of twice as many guess-and-correct cycles as the comparison group. An even more serious concern also arose: How do we know that learners in the comparison group actually engaged in comparison as part of their categorization process? The participants in this study knew that both members of each pair belonged to the same category, so if they felt able to classify one item (or if they only felt like considering one item), they were free to entirely ignore the other. In light of these limitations and challenges, our goal was to design an experiment that would satisfactorily test the power of comparison in learning relational categories via the classification paradigm.

Experiment

In order to investigate comparison while ensuring consideration of both items, we used a mix of within-category and cross-category pairs—participants made separate classification judgments for each item in each pair. Since the co-presented items might or might not belong to the same category, the learner is required to give direct consideration to both. Although this task still does not definitively require comparison, making two classification decisions on each trial gives rise to a subtle dynamic—the learner has to decide whether or not to guess the same category for each item. An implicit same/different category judgment is built into the explicit classification task. To reinforce this joint consideration, corrective feedback is provided only after both responses are made. Another important feature of this design is that comparison learners experience an equal number of item exposures and an equal number of guess-and-response cycles relative to the single-item control.

We also included a far-transfer task consisting of classification judgments (without feedback) using the same category labels in a completely different—though analogous—domain. Since the new materials (described below) had no surface similarity to the rock arrangements, this provides a particularly powerful test of relational learning. The far-transfer task addresses whether comparison-based learning leads to abstract schemas that facilitate generalization of the relational structure beyond the training domain.

Our core prediction was that comparison learners would more accurately classify old and new items relative to single-item learners. We expected the far-transfer task to show that comparison can promote on-the-spot generalization to a novel analogous domain and also to confirm the relational nature of the learning task.

Method

Participants. A total of 100 undergraduate students at Binghamton University participated for course credit.

Materials. The 36 stimulus items depicted unique rock arrangements consisting of from four to eight individual rocks of varied shape, size, and color. An arbitrarily selected, fixed subset of 24 examples was used for training (see Appendix), and the remaining items were used as transfer items. The rock arrangements represented three relational categories labeled “Tolar,” “Besod,” and “Makif.” Tolars were defined by the presence of two vertically stacked rocks, each of the same color and the same general shape. Besods were defined by the presence of one rock supported by two others. Makifs were defined by monotonically decreasing height from the left to the right of the arrangement. Each example conformed to exactly one of the relational categories (see Figure 1 for perceptual characteristics of the rock arrangements). In the comparison condition, an arbitrarily selected, fixed set of item pairings was used for all participants. There were equal numbers of within-category and cross-category pairs. An additional set of materials was developed to assess far transfer. These consisted of three categories of five “mobiles” (columns of colored geometric shapes connected by vertical line segments) corresponding to the underlying spatial structure of the rock arrangements (see Figure 1).

A potential concern was whether the categories could be learned based on some type of low-level perceptual similarity (such as similarity in global shape, size, or color) rather than the defining relational property—the perceptual variation among category members can be seen in the full set of training items (see Appendix). A separate study was conducted in order to validate the training materials for our purposes. The specific question was whether the task truly assesses relational learning or if there could be other stimulus characteristics (unbeknownst to the experimenters) that could lead to successful differentiation of the categories. We also sought to determine whether participants learn these particular categories through a strategy of memorizing individual exemplars or by developing abstract semantic representations. We tested 28 participants who learned to classify the rock arrangements into three categories (in accord with the main experiment below). Prior to categorizing each rock arrangement, participants were asked to describe what they noticed about the way the rocks were arranged. After typing their description into a response box on the computer screen, participants made a classification response. Initially, participants tended to vary their descriptions from

trial to trial. Over the course of the task, most participants settled upon and consistently applied a particular description for each category. We sought to identify a point in each participant's data at which their descriptions switched from perceptual characteristics (color, size, and number of rocks) to relational properties (positional and spatial arrangements consistent with the intended category definitions—sometimes with the use of terms like *cave*, *tower*, or *bridge*). On average, a total of 12.75 different descriptions ($SE = 1.31$) were used, and a relational shift occurred after 17.25 trials ($SE = 3.11$). Four participants (proportion of .14) failed to make a relational shift—these participants performed just above chance ($M = .46$, $SE = .03$) but scored significantly lower than the rest of the group ($M = .87$, $SE = .02$) on classification accuracy, $t(26) = 7.99$, $p < .001$. To assess the meaningfulness of the description shift in terms of classification performance, we computed mean accuracy for trials after the shift. In all cases, the switch to relational descriptions was accompanied by accuracy above 90% from that point forward. Further, participants who shifted earlier tended to show overall higher accuracy in the experiment ($r = -.70$, $p < .01$). These results provide good evidence that the experimental materials are consistent with the goal of assessing relational learning.

Procedure. Using a between-subjects design, we randomly assigned participants to either the comparison ($n = 50$) or single-item ($n = 50$) learning condition. All participants received an archaeology cover story including instructions to try to learn to tell which rock arrangements belonged to each of the three types. In the control condition, each learning trial began with a single training instance that remained on screen for the full trial. After the classification query, corrective feedback (evaluating whether right or wrong and providing the category label) was provided for a fixed interval of 3 s. The learning phase consisted of two passes through the training set of 24 items. Comparison with remembered examples cannot be avoided, so some degree of comparison-driven relational learning was expected in the control group. We were able to eliminate same-category comparison opportunities across consecutive trials by using a pseudorandom presentation order (the training item on each trial always belonged to a different category than that of the previous trial). Data from the preliminary study suggested that participants do not rely on this regularity—discovering and exploiting the regularity would be marked by rarely or never guessing the same category as the correct answer from the previous item. It does remain possible that there is some subtle or implicit learning of sequential structure (see Jones & Sieck, 2003) during the training phase.

In the comparison condition, participants were asked to classify one of the two co-presented instances and then the other. The position (left or right) of each image was determined randomly with alternation of whether the left or right item was queried first. After both responses were collected, corrective feedback for each item was given at the same time. The feedback appeared for a total of 6 s since there were two separate pieces of information to evaluate (resulting in an overall amount of time spent processing feedback equal to that of the single-item learners).

Given prior evidence that mere juxtaposition without an explicit comparative task can fail to elicit comparison (Catrambone & Holyoak, 1989; Kurtz et al., 2001; Loewenstein et al., 1999), we included a simple orienting task. Comparison learners were instructed, "Study the examples, then focus on a single rock in one

of the examples and consider the role it plays in that arrangement. Try to decide which rock plays a corresponding role in the other example." Participants made yes/no judgments as to whether the orienting task was "helpful." The orienting task for the single-item condition did not invite comparison: "Study the example, then focus on a single rock and consider the role it plays in the arrangement." We elected to maximize similarity of the orienting tasks despite the possibility that asking single-item learners to consider the "role" of a rock could promote relational encoding or temporal comparison across trials.

The testing procedure was identical for both conditions. Participants were presented with the 24 training items plus 12 new transfer items for classification in a random, intermixed order without feedback. For the test of far transfer, mobiles were shown one at a time in random order and learners were asked to classify them according to the same three category labels. We note that a set of pairwise similarity ratings for the rock arrangements was collected (as part of another line of inquiry) before the far transfer test phase.

Results and Discussion

Learning phase. While our key predictions are about the test measures, we begin by considering performance during the learning task. Participants in the single ($M = .63$, $SE = .02$) and comparison ($M = .59$, $SE = .03$) conditions did not differ in their accuracy in classifying the rock arrangements (recall that chance performance is .33). In a time-course analysis of classification performance, we found that comparison learners got off to a slow start ($M = .48$, $SE = .04$) compared to single-item learners ($M = .54$, $SE = .03$) in mean accuracy for the first third of the learning task but caught up by the final third: single ($M = .70$, $SE = .03$) and comparison ($M = .69$, $SE = .04$). This slow start was likely due to the uncertainty inherent in the nature of the comparison task. For example, the task of making two category guesses on a given trial may have created an initial bias toward guessing that the examples belonged to different categories. Such a response bias would likely be apparent early in learning and then be corrected with increasing category understanding. Consistent with this possibility, comparison learners were more likely to make different category guesses during the first third of the learning phase ($M = .69$, $SE = .02$) than during the rest of the learning phase ($M = .60$, $SE = .03$), $t(49) = 3.95$, $p < .01$, $d = 0.79$.

Test phase. Comparison learners ($M = .75$, $SE = .03$) were significantly more accurate at test than single-item learners ($M = .65$, $SE = .03$) on old items, $t(98) = 2.17$, $p < .05$, $d = 0.44$. For never-before-seen category members, comparison learners ($M = .72$, $SE = .03$) were again reliably more accurate than single-item learners ($M = .59$, $SE = .04$), $t(98) = 2.70$, $p < .01$, $d = 0.55$. Paired classification learning clearly promoted relational category acquisition. These results may in fact underestimate the effect of comparison since the transfer-appropriate processing framework (Morris, Bransford, & Franks, 1977) predicts an advantage for the control group based on the test task matching the training task.

There are several factors that can explain why a comparison advantage occurs at test but not during training (specifically with regard to the training items). The first thing to consider is the slow start for the comparison group due to task factors as noted above. Perhaps more important, given the relatively short learning phase

it is likely that a substantial subgroup of comparison learners achieved their relational insight at a point late in the learning phase—therefore, their category understanding was reflected only in the tail end of their learning performance but was fully reflected at test. Another possible factor is that participants understood the test phase as a chance to demonstrate their knowledge (and the learning phase as a chance to explore and improve), so participants may have held a clearer objective of achieving high accuracy during the test phase.

We used the far transfer test to evaluate whether comparison promotes not only acquisition of relational understanding in a domain but also further generalization of relational knowledge to a novel domain. No significant difference was seen between groups in the far transfer test, but this comes with an important caveat: Participants who did not learn well and did not grasp the source domain had no chance to achieve far transfer. We expected that an advantage for the comparison group might have been washed out by a group of unsuccessful learners performing at floor on far transfer. One way to address this issue is through an analysis of covariance controlling for variation in learning accuracy. Our goal in conducting such an analysis was to determine whether the study conditions affected performance on the far transfer task while correcting for variability in learning performance. We found a main effect of the learning condition on far transfer performance, $F(1, 97) = 4.28, p < .05$, with participants in the comparison condition ($M = .67, SE = .03$) outperforming those in the single condition ($M = .57, SE = .03$).

We conducted an additional analysis using a criterion for inclusion based on mean accuracy in the initial learning phase. We defined “good” learners as those who attained learning accuracy of 60% or better (30 participants in the single condition and 26 participants in the comparison condition), and the remainder were classified as “bad” learners (20 participants in the single condition and 24 participants in the comparison condition). A majority of participants were assigned to the “good” learner status, but the sizable minorities of participants in both conditions who struggled to master the categories were excluded. For verification purposes, we also conducted the analysis using a criterion for inclusion based on accuracy within one standard deviation of the overall mean; this led to consistent outcomes on statistical tests.

Looking at learning performance, we found no difference in mean accuracy between the single and comparison conditions for either the good learners (single: $M = .73, SE = .02$; comparison: $M = .74, SE = .02$) or the bad learners (single: $M = .48, SE = .02$; comparison: $M = .42, SE = .02$). We note that the average accuracy in training performance for the bad learners was close to the border for a significant difference from chance (.46 according to binomial distribution), meaning that it is questionable whether these participants had learned anything at all. We observed a marginally significant difference between conditions during the last third of the learning phase when looking at the performance of the good learners (single: $M = .80, SE = .02$; comparison: $M = .87, SE = .03$), $t(54) = 1.98, p = .053$. This suggests that following initial difficulty at the start of the learning phase, good learners were better able to master the categories when given the advantage of side-by-side comparison. The observed advantages for the comparison group in classifying old and new rock arrangements at test were also maintained for the good learners.

Of primary concern in this analysis, we observed a comparison advantage in far transfer among participants who had learned the source domain well enough to support such transfer. The comparison group ($M = .78; SE = .05$) was significantly more accurate than the single group ($M = .65; SE = .04$) on mean accuracy in far transfer, $t(54) = 2.15, p < .05$. These far transfer results in conjunction with the significant difference in the analysis of covariance nicely underscore the comparison advantage: Comparison experience leads to better generalization of category knowledge to a novel, analogous domain in which there is an entirely distinct surface manifestation of the underlying relations.

An additional implication of the far transfer results is further clarification (along with the validation study using the description task) that learners acquired the relational basis for the categories, as opposed to picking up on idiosyncratic perceptual properties or memorizing category associations. We found that both groups performed well above chance, $p < .001$, on the far transfer test. It is not clear how learners could have achieved any level of success at all on this task without having picked up on the relational content underlying both the rock arrangements and the mobiles.

General Discussion

The goal in these studies was to test the prediction that side-by-side comparison promotes relational category learning. A preliminary study evaluated a straightforward extension of the classification learning paradigm to include comparison opportunities but failed to demand comparison in the task of making joint classification decisions for within-category pairs.

The main experiment required explicit consideration of each co-presented item and implicitly tasked the learner with engaging in joint evaluation. At test, the comparison group performed better than the single-item control group on old items and on tests of within-domain and cross-domain transfer.

While there is much that might be done to further this research, we have shown that properly constructed comparison opportunities promote relational category learning. This finding adds to the available evidence on the power of comparison by demonstrating how embedding comparison opportunities in a classification learning task promotes relational learning. While structure-mapping theory predicts that same-category pairs should promote relational abstraction more than cross-category pairs, our mixed-pairs design enforces consideration of both items and creates an implicit same/different task such that the comparison and classification components become closely integrated. The cross-category pairs also provide a potentially beneficial opportunity for learning by contrast. One promising direction for future work is to link these findings in relational category learning to research on comparison of exemplars in the attribute-based category learning literature (Andrews, Livingston, & Kurtz, 2011; Hammer, Hertz, Hochstein, & Weinshall, 2009; Helie & Ashby, 2012; Higgins & Ross, 2011; Spalding & Ross, 1994).

Our focus on relational categories and the comparison of co-presented training examples accords with increasing awareness of the importance of relational content in the processing of real-world categories (Gentner & Kurtz, 2005; Markman & Stillwell, 2001; Medin, Goldstone, & Gentner, 1993; Murphy & Medin, 1985; Schyns, Goldstone, & Thibaut, 1998). Kloos and Sloutsky (2008) found that relationally defined categories were learned much more

effectively through direct instruction than through observational learning without feedback. Perhaps the use of supervised classification plus comparison opportunities would be a stronger competitor with direct instruction for ease of acquisition and might prove to be a more effective basis for transfer. Since promoting spontaneous transfer of learned relational concepts to novel settings is one of the great challenges for relating psychological research to instructional practice (Barnett & Ceci, 2002), we hope to build on the present evidence that comparison-based learning and classification learning can be integrated into a successful technique for acquiring and generalizing relational categories.

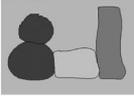
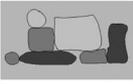
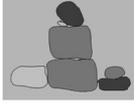
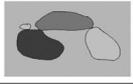
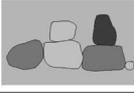
References

- Andrews, J. K., Livingston, K. R., & Kurtz, K. J. (2011). Category learning in the context of co-presented items. *Cognitive Processing, 12*, 161–175. doi:10.1007/s10339-010-0377-5
- Asmuth, J., & Gentner, D. (2005). Context sensitivity of relational nouns. In *Proceedings of the Twenty-seventh Annual Meeting of the Cognitive Science Society* (pp. 163–168). Mahwah, NJ: Erlbaum.
- Asmuth, J., & Gentner, D. (2013). *Relational categories are more mutable than entity categories*. Manuscript in preparation.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612–637. doi:10.1037/0033-2909.128.4.612
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 1147–1156. doi:10.1037/0278-7393.15.6.1147
- Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373. doi:10.1080/15248371003700015
- 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. doi:10.1037/0033-295X.115.1.1
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170. doi:10.1207/s15516709cog0702_3
- Gentner, D. (1989). Mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). London, England: Cambridge University Press.
- Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 245–275). Hillsdale, NJ: Erlbaum.
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*, 752–775. doi:10.1111/j.1551-6709.2010.01114.x
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure-mapping and relational language support children's learning of relational categories. *Child Development, 82*, 1173–1188. doi:10.1111/j.1467-8624.2011.01599.x
- Gentner, D., & Asmuth, J. (2008). Can relationality be distinguished from abstractness in noun mutability? In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 863–868). Austin, TX: Cognitive Science Society.
- Gentner, D., Holyoak, K. J., & Kokinov, B. (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.
- Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151–175). Washington, DC: American Psychological Association.
- Gentner, D., Levine, S., Dhillon, S., & Poltermann, A. (2009). Using structural alignment to facilitate learning of spatial concepts in an informal setting. In B. Kokinov, K. Holyoak, & D. Gentner (Eds.), *Proceedings of the Second International Conference on Analogy*. Sofia, Bulgaria: NBU Press.
- Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. D. (2009). Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science, 33*(8), 1343–1382. doi:10.1111/j.1551-6709.2009.01070.x
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56. doi:10.1037/0003-066X.52.1.45
- Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*, 263–297. doi:10.1016/S0010-0277(98)00002-X
- Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513. doi:10.1016/S0885-2014(99)00016-7
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38. doi:10.1016/0010-0285(83)90002-6
- Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin & Review, 18*, 406–413. doi:10.3758/s13423-011-0058-0
- Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition, 118*, 359–376. doi:10.1016/j.cognition.2010.10.009
- Hammer, R., Hertz, T., Hochstein, S., & Weinshall, D. (2009). Category learning from equivalence constraints. *Cognitive Processing, 10*, 211–232. doi:10.1007/s10339-008-0243-x
- Helie, S., & Ashby, F. G. (2012). Learning and transfer of category knowledge in an indirect categorization task. *Psychological Research, 76*, 292–302. doi:10.1007/s00426-011-0348-1
- Higgins, E. J., & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1388–1393). Austin, TX: Cognitive Science Society.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466. doi:10.1037/0033-295X.104.3.427
- Jones, M., & Love, B. C. (2007). Beyond common features: The role of roles in determining similarity. *Cognitive Psychology, 55*, 196–231. doi:10.1016/j.cogpsych.2006.09.004
- Jones, M., & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 626–640. doi:10.1037/0278-7393.29.4.626
- Kloos, H., & Sloutsky, V. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General, 137*, 52–72. doi:10.1037/0096-3445.137.1.52
- Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822. doi:10.2307/1131753
- Kurtz, K. J. (2007). The divergent autoencoder (DIVA) model of category learning. *Psychonomic Bulletin & Review, 14*, 560–576. doi:10.3758/BF03196806
- Kurtz, K. J., & Gentner, D. (1998). Category learning and comparison in the evolution of similarity structure. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (p. 1236). Mahwah, NJ: Erlbaum.
- Kurtz, K. J., & Gentner, D. (2001). Kinds of kinds: Sources of category coherence. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 522–527). Mahwah, NJ: Erlbaum.

- Kurtz, K. J., & Gentner, D. (2013). *Putting kinds into kinds: On the nature of second-order categories*. Manuscript in preparation.
- Kurtz, K. J., & Loewenstein, J. L. (2007). Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one. *Memory & Cognition*, *35*, 334–341. doi:10.3758/BF03193454
- Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences*, *10*, 417–446. doi:10.1207/S15327809JLS1004new_2
- Levering, K., & Kurtz, K. J. (2010). Generalization in higher-order cognition: Categorization and analogy as bridges to stored knowledge. In M. T. Banich & D. Caccamisse (Eds.), *Generalization of knowledge: Multidisciplinary perspectives* (pp. 175–196). New York, NY: Psychology Press.
- Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review*, *6*, 586–597. doi:10.3758/BF03212967
- Markman, A. B., & Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence*, *13*, 329–358. doi:10.1080/09528130110100252
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278. doi:10.1037/0033-295X.100.2.254
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. doi:10.1016/S0022-5371(77)80016-9
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316. doi:10.1037/0033-295X.92.3.289
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, *131*, 5–15.
- Pothenos, E. M., & Wills, A. J. (Eds.). (2011). *Formal approaches in categorization*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511921322
- Rehder, B., & Ross, B. (2001). Abstract coherent categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1261–1275. doi:10.1037/0278-7393.27.5.1261
- Rein, J. R., Goldwater, M. B., & Markman, A. B. (2010). What is typical about the typicality effect in category-based induction? *Memory & Cognition*, *38*, 377–388. doi:10.3758/MC.38.3.377
- Rittle-Johnson, B., & Star, J. (2009). Compared to what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*, 529–544. doi:10.1037/a0014224
- Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 42–55. doi:10.1037/0278-7393.16.1.42
- Ross, B., Perkins, S., & Tenpenny, P. (1990). Reminding-based category learning. *Cognitive Psychology*, *22*, 460–492. doi:10.1016/0010-0285(90)90010-2
- Ross, B. H., Taylor, E. G., Middleton, E. L., & Nokes, T. J. (2008). Concept and category learning in humans. In J. Byrne (Series Ed.) & H. L. Roediger III (Vol. Ed.), *Learning and memory: A comprehensive reference: Vol 2. Cognitive psychology of memory* (pp. 535–557). Oxford, England: Elsevier.
- Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences*, *21*, 1–17. doi:10.1017/S0140525X98000107
- Spalding, T. L., & Ross, B. (1994). Comparison-based learning: Effects of comparing instances during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1251–1263. doi:10.1037/0278-7393.20.6.1251
- Tomlinson, M., & Love, B. (2010). When learning to classify by relations is easier than by features. *Thinking & Reasoning*, *16*, 372–401. doi:10.1080/13546783.2010.530464
- Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*, 719–736. doi:10.1207/s15516709cog0000_33

(Appendix follows)

Appendix
All Rock Arrangements in Learning Phase

Category		
Besod	Makif	Tolar
		
		
		
		
		
		
		
		

Received June 13, 2012
Revision received October 1, 2012
Accepted October 2, 2012 ■