

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

The Role of Difference-Detection in Learning Contrastive Categories

#### **Permalink**

<https://escholarship.org/uc/item/7105k5q0>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

#### **ISSN**

1069-7977

#### **Authors**

Smith, Linsey  
Gentner, Dedre

#### **Publication Date**

2014

Peer reviewed

# The Role of Difference-Detection in Learning Contrastive Categories

Linsey A. Smith (linsey@u.northwestern.edu)

Department of Psychology, 2029 Sheridan Road  
Evanston, IL 60208 USA

Dedre Gentner (gentner@northwestern.edu)

Department of Psychology, 2029 Sheridan Road  
Evanston, IL 60208 USA

## Abstract

Prior research has found that comparison fosters abstraction and transfer of concepts (e.g., categories, solution methods). These learning benefits are often explained by virtue of comparison's ability to highlight common relational structure between cases. Here we explore the role of comparison in identifying critical differences. Participants compared contrastive cases, listed differences between them, and completed a classification task. We found that carrying out a structural alignment prior to listing differences influenced the kinds of differences people noticed. Further, the kinds of differences people noticed predicted their subsequent classification performance.

**Keywords:** Analogy; structural alignment; comparison; contrast; learning

## Introduction

Comparison has been shown to lead to learning in a number of different realms for both children and adults. Comparing cases facilitates transfer and problem-solving in adults (e.g., Catrambone & Holyoak, 1989; Gentner, Loewenstein, Thompson, & Forbus, 2009; Gick & Holyoak, 1983). Comparison also fosters children's learning of relational categories (Gentner, Anggoro & Klibanoff, 2011) and relational language (Childers, 2011; Gentner & Namy, 2006; Haryu, Imai & Uchida, 2011). A recent meta-analysis by Alfieri et al., (2013) found that the use of comparison in classrooms is a strong predictor of learning gains.

How do these benefits come about? According to structure-mapping theory (Gentner, 1983), when learners compare two cases, they generate a structural alignment between the two representations. This fosters learning in at least three ways (Gentner, 2010; Gentner & Markman, 1997). First, it increases the salience of their common structure; second, it invites inferences from one case to the other; and, third, it highlights *alignable differences*—differences connected to the common structure.

Much of the research showing positive effects of comparison on learning has focused on its effects in abstracting commonalities and inviting inferences (e.g., Catrambone & Holyoak, 1989; Gentner et al., 2009). However, there is mounting evidence that comparison can aid in differentiation as well as in abstraction. For example, comparing two “near-miss” cases (McClure, Friedman, & Forbus, 2010), which are identical except for a crucial structural difference, improves learning (e.g., Gick &

Paterson, 1992). Comparison also fosters discrimination between more complex cases, such as alternative solution methods (Rittle-Johnson & Star, 2009), easily confusable concepts (e.g., Day, Goldstone & Hills, 2010; VanderStoep & Seifert, 1993), and category exemplars vs. non-exemplars (Gick & Paterson, 1992; Kok, de Bruin, Robben, & van Merriënboer, 2012; Kurtz & Gentner, 2013). For example, Day et al. (2010) found that having middle-school students contrast positive and negative feedback systems could improve classification of new examples. An open question is how exactly the observed learning effects come about in contrastive case comparisons. These findings underscore the benefit of *contrastive cases* in learning. Many of these studies utilize pairs that are highly similar except for the crucial difference. Such pairs have two advantages. First, they are ‘self-aligning’—that is, they are extremely easy to align, even for children and novices. Second, once aligned, they have few or no competing alignable differences besides the key intended difference. For example, Kurtz and Gentner (2013) found that people could identify an error in a skeleton faster if they compared it with a highly alignable correct example than if they compared it with the same correct example mirror-reversed (and thus less perceptually alignable).

But not all important distinctions can be illustrated with very close ‘near-miss’ pairs. Many important category distinctions involve moderate similarity, with some overlap and many differences. Here we ask what kinds of learning processes best facilitate learning in these more complex cases, in which pairs from different categories are only moderately similar—not so close as to be “self-aligning.” Because structural alignment highlights not only commonalities but also alignable differences, we propose that explicitly encouraging comparison between members of the two categories will facilitate noticing differences and thereby facilitate learning the category distinction.

Prior work has shown a relationship between structural alignment and difference-detection (e.g., Gentner & Gunn, 2001; Sagi et al., 2012) and between comparing contrastive cases and transfer (e.g., Day, Goldstone, & Hills, 2010; Rittle-Johnson & Star, 2009); our goal is to clarify the relationships between these phenomena. Thus, the current study examines the connections between structural alignment, difference-detection, and subsequent ability to classify new examples of the two categories.

We chose positive and negative feedback systems as our

domain of inquiry. These two kinds of systems are of wide importance across both physical and social domains, yet prior work suggests that these causal structures are not salient, even to college students (Rottmann, Gentner, & Goldwater, 2012; see also Day et al., 2010). These categories make an ideal testing ground because they share significant common structure as well as significant differences. As feedback systems, they are both causal systems that share the commonality that both are causal systems in which part of the output is returned to serve as input to the same system. Thus in both cases the output ultimately re-affects itself. These two kinds of systems also differ in a fundamental way: In *positive feedback* the output increases the input; this in turn produces a greater output. This results in a cycle of increasing magnitude of effect. In *negative feedback*, the output reduces the input. Thus if the input is increased, this will increase the output, which will decrease the subsequent input. This results in a cycle that stabilizes the system.

In our study, participants were given two contrastive cases—a positive feedback case and a negative feedback case, exemplified in Table 1. Although the two cases were always from the same domain, they differed in many ways beyond the positive-vs.-negative feedback distinction. For example, the two cases in Table 1 involve different hormones, and they differ in that one concerns healthy cells and the other, cancer cells. Participants were told that one of the cases was an example of “System A” and the other, of “System B.” Participants listed differences between them and went on to complete a transfer task in which they had to classify examples into these two systems. The key manipulation was that half the participants were asked to list commonalities before listing differences. Then participants went on to carry out a classification task in which they had to decide whether new phenomena were exemplars of System A or System B—that is, of positive vs. negative feedback.

We assume, based on prior research, that listing commonalities will induce people to carry out a structural alignment between the two cases, and that this will lead them to focus on the maximal common relational structure—namely, that both are causal systems and that in both of them, the output is ‘fed back’ into the system. We hypothesize that structurally aligning the two cases will lead people to notice alignable differences connected to this causal structure—leading to greater likelihood of noticing the key difference between positive and negative feedback. We further predict that people who have noticed a key difference will be better able to classify further phenomena as positive versus negative feedback systems than those who have not.

Thus, our hypotheses are (1) that people who list commonalities before they list differences will be more likely to produce key differences than those who simply list differences; and (2) that listing key differences between the study pair of exemplars will be causally related to better performance on the classification task.

The first hypothesis—that people who list commonalities

and who therefore carry out a structural alignment between the cases) will be more likely to notice the key alignable difference than those who do not—has some support from prior findings. Gentner and Gunn (2001) found that people were more likely to list a difference (typically, an alignable difference) between a pair of concepts if they had previously listed commonalities for that pair. More to the point, Kurtz, Miao, & Gentner (2001) gave people two physics scenarios and varied the level of comparison intensity that participants were asked to engage in—from separate descriptions (no comparison) to similarity ratings (shallow comparison) to writing out commonalities and stating correspondences (intensive comparison). Then participants were given both scenarios and asked to state a difference between them. People were more likely to list an important alignable difference—that is, one that was related to the key causal commonality (heat flow)—if they had engaged in more intensive comparison. These results lend support for the prediction that inducing participants to list commonalities will foster noticing key alignable differences.

However our second hypothesis—that participants who list key differences will perform better on the classification task—needs some unpacking. A link between difference detection and later classification performance could come about in at least two ways. One possibility—the one we favor—is that detecting a key difference is *causally related* to learning and transfer to the classification task. But an alternate possibility is that identifying a key difference is simply *diagnostic* of having achieved a structural alignment—better alignment (as prompted by a commonality task) leads both to better difference-detection and to better transfer. To this end, we test whether difference-detection *mediates* the link between structural alignment (commonality listing) and classification task performance. Finding this mediation effect will provide strong support for a causal link between difference-detection and discrimination learning.

## Experiment

Participants were presented with two contrastive cases, listed any differences they noticed between them, then completed a classification task where they had to identify new cases as either examples of positive or negative feedback.

### Participants

152 participants were recruited through Amazon Mechanical Turk. Participants were all located in the United States and appeared to be fluent English speakers. The task required 15–20 minutes.

### Materials and Design

**Difference Listing Task** Each participant was randomly assigned to one of two conditions: Explicit Comparison or Control. In the Control condition, participants were asked to “Write a difference between System A and System B (feel free to list more than one if you choose).” In the Explicit Comparison condition, participants were invited to compare

the cases (“Write out any important similarities between System A and System B”) before they listed differences. The experiment was presented within the Qualtrics interface.

**Contrastive Cases** Examples of positive and negative feedback were drawn from four domains: ecology, economics, engineering, and physiology. The two feedback cases from physiology are shown in Table 1.

Table 1. Sample positive (System A) and negative (System B) feedback passages, from the physiology domain

<b>System A</b>
<i>When a tumor grows, cancerous cells manufacture excess growth hormones. These hormones send a signal that causes the body to increase production of cancer cells. This increase in cancer cells leads to a further increase in growth hormones. This in turn leads the body to produce even more cancer cells, which leads to an even higher concentration of growth hormones. The further increase in growth hormones encourages additional proliferation of cancer cells.</i>
<b>System B</b>
<i>In normal cell proliferation, healthy cells release growth inhibitory factors—chemicals that tell the body to reduce the rate of cell growth. As the body produces more cells, the concentration of growth inhibitory chemicals in the body increases. Consequently, the body begins to produce fewer new cells. With fewer new cells being produced, there is a reduction in the amount of inhibitory chemicals produced. Once again, cell production begins to increase.</i>

Participants received two examples from a single domain. To ensure that any effects we observed were tied to the actual case comparisons rather than to knowledge of the terms *positive* and *negative feedback*, we introduced each case using the generic labels *System A* and *System B*. The selection of cases was randomized and their position on the screen was counterbalanced.

**Classification Task** The classification task was designed to assess learners’ ability to discriminate between positive and negative feedback systems. These materials included 16 brief scenarios, each describing a real-world phenomenon (some scenarios were adapted from Rottmann et al., 2012). Eight of these scenarios described positive feedback systems, and eight described negative feedback systems. Furthermore, the 16 scenarios were drawn from the same four domains that were used in the contrastive cases—this allowed us to assess how narrow or broad peoples’ category generalizations were. Feedback type and Domain varied orthogonally—there were two examples of positive feedback and two examples of negative feedback for each of the four domains. Participants classified each scenario as an example of either a positive or negative feedback system by selecting one of four multiple choice options: *System A*, *System B*, *neither*, or *I don’t know*. The scenarios were presented in

random order. Because the 16 scenarios were drawn from the same four domains that had been used in the study cases, each participant saw four examples from the studied domain and twelve from non-studied domains. This allowed us to assess how narrow or broad peoples’ category generalizations were.

## Measures

**Difference Ratings** We coded several aspects of participants’ difference listings. First, we coded whether or not the participant provided a *key difference*—the crucial alignable difference that captures the distinction between positive and negative feedbacks. For key differences, we further coded them as being: (1) a *global* or *process-level* description, and (2) an *abstract* or *concrete* description. A *global key difference* captures the higher-order difference between positive and negative feedbacks. A *process-level key difference* invokes a distinction in the specific causal structures of the cases. Participants who gave process-level descriptions typically used language expressing qualitative relations such as “X increased [decreased] Y.” A *concrete* key difference describes the difference by referring to particular elements in the passages (e.g., cells), whereas an abstract key difference does not. The first two examples below exemplify global descriptions at the concrete and abstract levels.

*System B controls the rate at which cells grow. In System A, the cancerous cells have the ability to grow with no inhibition, unlike System B which keeps itself balanced out. (Global, Concrete)*

*The difference between System A and B is that in System A, one action causes a reaction that eventually leads to a balance. In System B, the reaction continues to increase in intensity until it is out of control. (Global, Abstract)*

The next two examples illustrate process-level descriptions at the concrete and abstract levels.

*System A just produces more cancer cells as growth hormones increase. System B regulates the production of cells producing less with the increase in inhibitory chemicals and more when there is a reduction of inhibitory chemicals. (Process, Concrete)*

*System B describes a system that increases, then decreases, then increases as production increases and decreases, whereas System A describes a system that only increases no matter the production. (Process, Abstract)*

Two raters, blind to instructional condition, coded the differences. Inter-rater Reliability was high for all measures (Key Difference:  $\kappa=.87$ ,  $p<.001$ ; Global/Process:  $\kappa=.88$ ,  $p<.001$ ; Abstract/Concrete:  $\kappa=.90$ ,  $p<.001$ ). Disagreements were resolved through discussion.

**Classification Score** Each participant received a score based on total correct answers (max=16). To assess the broadness of transfer, we calculated separate scores for scenarios drawn from the training domain (Close Transfer Score) and domains that the participant was not trained on (Far Transfer Score). For example, if an individual compared cases from Engineering, their Close Transfer Score consisted of their total score on the four scenarios from Engineering (max=4), whereas their Far Transfer Score would consist of their score on the scenarios from Economics, Ecology, and Physiology (max=12). All scores were converted to percent correct.

**Predictions** There are two variables of interest here: key difference detection and classification task performance. Compressing our hypotheses into the key predictions, we predict, first, that listing similarities will induce a structural alignment; thus people who explicitly list similarities before differences will more often produce the key alignable difference than people who only list differences; and, second, that key difference detection will lead to better understanding of the distinction between positive and negative feedback, and therefore to better performance on the classification task. A strong version of this second prediction—that difference detection is causally related to discrimination learning—further predicts that key difference-detection should mediate the relationship between Explicit Comparison and Classification Score.

**Results**

To test our first prediction—that similarity-listing will induce a structural alignment and that this will foster key difference-detection—we examined how Explicit Comparison influenced key difference identification (Figure 1). Overall, people who listed similarities produced the key difference 57% of the time versus 28% of the time for people who only listed differences. A binary logistic regression revealed that Task (Explicit Comparison vs. Control) was a significant predictor of whether people produced the key difference. People who listed similarities were more likely to generate the key difference than people who did not list similarities (Wald  $\chi^2 = 12.73$ ,  $df=1$ ,  $p<.001$ , Odds Ratio=3.42).

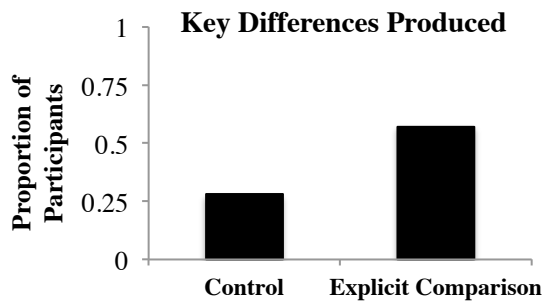


Figure 1: Key differences produced, by Task

To test our second hypothesis—that difference-detection predicts performance on the classification task—we assessed

the relationship between key difference detection and later classification. Figure 2 shows the dispersion of scores for individuals who produced the key difference and those who did not. Participants who detected the key difference had higher scores on the classification test ( $m=0.79$ ,  $SD=0.16$ ) than participants who did not identify the key difference ( $m=0.52$ ,  $SD=0.23$ ),  $t(149.89)=8.61$ ,  $p<.001$ ,  $d=1.37$ . Exploratory analyses revealed that that the effect size was much larger for Far Transfer ( $d=1.50$ ) than for Near Transfer problems ( $d=0.88$ ), suggesting a particularly strong relationship between difference-detection and far transfer.

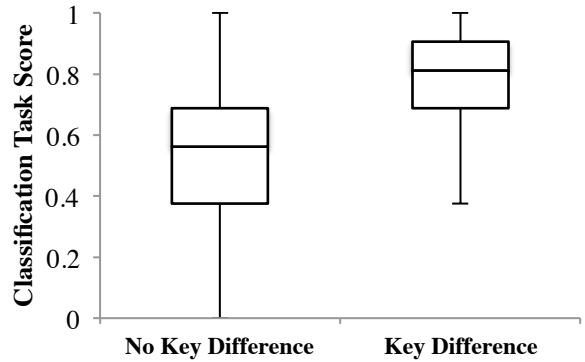


Figure 2: Dispersion of classification task scores, by key difference production

To test the plausibility of a causal link between difference-detection and discrimination—a strong interpretation of our second hypothesis—we conducted a simple mediation analysis (Hayes, 2013). This analysis involved a standard three-variable path model (Figure 3; Baron & Kenny, 1986). Three tests must reach statistical significance to conclude mediation: (1) the initial variable (Task) must be related to the mediating variable (Difference-Detection; Figure 3, Path a); (2) the mediating variable must be related to the outcome variable (Classification Score) after controlling for the initial variable (Figure 3, Path b); and (3) the mediation effect ( $a*b$ ) must be significant (Hayes, 2013).

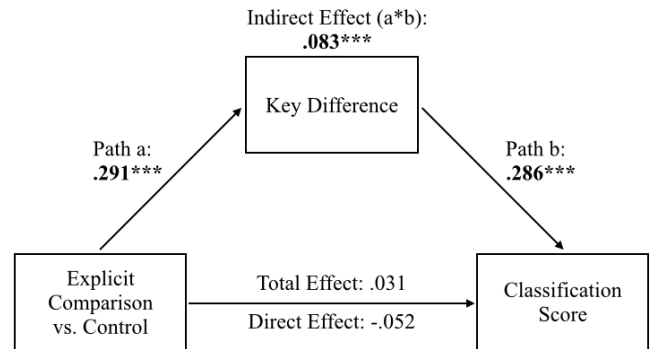


Figure 3: Mediation analysis. Unstandardized coefficients are shown, significant at \*\*\* $p < 0.001$ .

The mediation analysis indicates that explicit comparison indirectly influenced classification performance through its

effect on key difference detection. As can be seen in Figure 3, participants who listed similarities were more likely to notice the key difference ( $a=.29, p<.001$ ), and participants who noticed the key difference did better on the classification task ( $b=.286, p<.001$ ). To put this another way, participants who listed a key difference scored on average 28.6% above participants who did not list the key difference. The indirect effect was significant<sup>1</sup> ( $p<.001$ ). The direct effect was non-significant, indicating that explicit comparison did not influence classification task performance independent of its effect on key difference detection (direct  $=-.052, p>.05$ ).

**Types of Key Differences and Later Transfer** An examination of the kinds of key differences produced revealed that, overall, 54% of key differences were process-level descriptions and 46% were global descriptions, a non-significant difference,  $z=0.5, p=0.61$ . People also produced many more Concrete differences (83%) than Abstract differences (17%),  $z=5.04, p<.001$ . We explored two aspects of the data concerning the kinds of key differences produced. First, we asked whether there was a connection between explicit comparison and the types of key differences generated. For instance, listing similarities between cases may have lead to the production of more process-level differences since people explicitly specified the common relational structure. However, the distribution of difference types did not vary by task,  $\chi^2(3)=.63, p=0.89$ . There was no evidence that explicit comparison influenced the kinds of differences people generated here.

Second, we asked whether the type of key difference was related to breadth of transfer. For example, if a person generated an abstract difference they may hold a more general representation of the key category distinction; therefore they may do better on Far Transfer than a person who produced a concrete difference. To identify connections between the types of key differences produced and transfer, we ran three separate regression models for the dependent measures of Overall Transfer, Close Transfer, and Far Transfer. Explanation Type (Process/Global) and Abstractness (Abstract/Concrete) were entered as predictors. Explanation Type was a significant predictor of far transfer ( $\beta=0.32, p<.05$ )—people who produced global differences scored higher on far transfer than people who produced process-level differences. No other significant relationships between key difference type and transfer were found.

## Discussion

There are two main findings. First, we found evidence for our first hypothesis—that structural alignment would facilitate key difference-detection. People who explicitly compared cases and stated commonalities were more likely to go on to generate a key alignable difference between the

positive and negative feedback cases than were those who simply listed differences. This result is consistent with prior work that demonstrates (a) that structural alignment favors discovering the maximal common relational structure (Clement & Gentner, 1991)—in this case the causal structure of feedback systems—and (b) that structural alignment facilitates detection of alignable differences connected to the common structure (Markman & Gentner, 1993a; 1994; Gentner & Gunn, 2001).

In line with our second hypothesis—that difference detection predicts discrimination learning—we found that people who generated a key difference did better on the subsequent classification task than people who did not generate the key difference, suggesting that difference-detection plays a critical role in learning. Indeed, the data support a particularly strong version of the second hypothesis. We found that difference-detection mediated the link between structural alignment (or more precisely, between whether people received explicit comparison instructions) and classification performance. There was no direct effect of explicit comparison on later classification—that is, the effect of explicit comparison on contrastive case learning was only observed through its effect on key difference detection. This suggests that the effect of structural alignment here was to promote key difference-detection, which in turn led to higher performance on the classification task.

Prior research has found that comparison fosters abstraction and transfer of concepts (e.g., categories, solution methods). These learning benefits are often explained by virtue of comparison's ability to highlight common relational structure between cases. There has been far less attention to the role of structural alignment in identifying critical differences. (but see Hammer, Diesendruck, Weinshall, & Hochstein, 2009; Higgins & Ross, 2011; Rittle-Johnson & Star, 2009). Our finding that focusing on commonalities prior to listing differences led people to notice more critical differences—and that this improved subsequent classification performance—suggests ways of better using comparison processes to foster discrimination learning.

Of course, we would not want to claim that comparing two cases from different categories will necessarily lead to deeper understanding of the key category differences. We suspect that this effect depends on there being significant structural overlap between the categories, as is the case for positive and negative feedback. But because overlapping category descriptions are likely to be difficult to discriminate, the methods described here may have an important role to play in complex learning, including learning in mathematics and science. Future work should explore to what degree the patterns found here extend to other categorical contrasts.

There are specific limitations in the current study that our ongoing work aims to address. First, we are verifying that key-difference facilitation is due to structural alignment in particular rather than simply a consequence of spending more time processing the cases (since the Control group

---

<sup>1</sup>A bias-corrected bootstrap confidence interval for the indirect effect ( $a*b=.083$ ) based on 5,000 bootstrap samples was entirely above zero (.039 to 0.137,  $p=.001$ ).

didn't engage in an alternative task while the Explicit Comparison group listed commonalities). Second, while the mediation analysis provides strong support for a causal link between difference-detection and discrimination learning, convergent evidence is required to substantiate this claim. In current studies we are directly manipulating difference-detection to determine its effects on classification task performance.

There remain many open questions regarding the links between comparison, difference-detection, and learning. For example, how might other factors of the comparison task, such as surface similarity between the cases, influence difference-detection and learning? Perhaps key difference detection is best when the features of the compared cases are relatively similar to one another, versus when their content is more dissimilar. However, comparison of dissimilar cases may lead to key difference representations with broader generalizability, whereas comparison of highly similar contrastive cases may lead to relatively specific common representations, which could limit transfer (Goldstone & Sakamoto, 2003).

## Conclusions

We find that the simple act of comparing two contrastive cases is helpful in learning to distinguish between two complex interrelated relational concepts. These findings offer insight the role of comparison in learning from contrastive cases.

## Acknowledgments

This work was supported by Office of Naval Research grant N00014-13-1-0470. We would like to thank Christian Dayal for his help with coding.

## References

- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48*(2), 87-113
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *JEP: LMC, 15*(6), 1147-1156.
- Day, S., Goldstone, R. L., & Hills, T. (2010). The effects of similarity and individual differences on comparison and transfer. In S. Ohlsson, and R. Catrambone, (Eds.), *Cognition in Flux: Proc. of the 32nd annual meeting of the Cognitive Science Society* (pp. 465-470)..
- Forbus, K. (1984). Qualitative process theory. *Artificial Intelligence, 24*, 85-168.
- Forbus, K. D. (2003). Qualitative Reasoning. In *CRC Handbook of Computer Science and Engineering, 2<sup>nd</sup> Edition*. CRC Press.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.
- Gentner, D., & Gunn, V. (2001). Structural alignment facilitates the noticing of differences. *Memory and Cognition, 29*(4), 565-577.
- Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. (2009) Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science, 3*, 1343-1382.
- Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science, 5*(3), 152-158.
- Gentner, D., & Markman, A.B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45-56.
- Gentner, D. & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*, 297-301.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.
- Gick, M. L., & Paterson, K. J. (1992). Do contrasting examples facilitate schema acquisition and analogical transfer? *Canadian Journal of Psychology, 46*, 539-550.
- Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition, 112*(1), 105-119.
- Haryu, E., Imai, M., & Okada, H. (2011) Object similarity bootstraps young children to action-based verb extensions. *Child Development, 82*(2). 674-686.
- Higgins, E. J., & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Kok, E. M., de Bruin, A. B. H., Robben, S. G. F., & van Merriënboer, J. J. G. (2013). Learning radiological appearances of diseases: Does comparison help? *Learning and Instruction, 23*, 90 -97.
- Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *JEP: LMC, 39*(4), 1303-1310.
- Kurtz, K. J., & Gentner, D. (2013). Detecting anomalous features in complex stimuli: The role of structured comparison. *JEP: Applied, 19*(3), 219-232.
- Kurtz, K., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *J. Learning Sci., 10*, 417-446.
- Markman, A. B., & Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *J. Memory and Language, 32*, 517-535.
- McLure, M. D., Friedman, S. E., & Forbus, K. D. (2010). Learning concepts from sketches via analogical generalization and near-misses. *Proceedings of the 32nd annual meeting of the Cognitive Science Society* (pp. 465-470).
- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *J. Educational Psychology, 101*(3), 529-544.
- Sagi, E., Gentner, D. & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36*(6), 1019-1050.