Relations are slower to encode than objects, explaining the same-different paradox

Yinyuan Zheng[1,2] (zhengyinyuan@berkeley.edu)
Dedre Gentner[1] (gentner@northwester.edu)
[1]Department of Psychology, Northwestern University
2029 N. Sheridan Rd., Evanston, IL 60208, USA
[2]Department of Psychology, University of California, Berkeley
2121 Berkeley Way, Berkeley, CA 94704, USA

Abstract

Relations are key to everyday cognition, but research has suggested that reasoning about relations is more challenging than reasoning about objects. We propose that one source of the difficulty resides in encoding: relations are slower and more resource-intensive to encode than objects. Two experiments provide evidence for this claim. Participants completed two straightforward matching tasks: given a standard, they had to identify either a relational match from a nonmatch, or an object match from a nonmatch. The critical manipulation was how long the standard was presented (i.e., the time allowed for encoding). The results showed that as presentation duration diminished, accuracy for relational matching deteriorated steeply, while that for object matching was relatively unaffected. This supports the idea that relational encoding is more resource-intensive. This pattern held for both similarity (Experiment 1) and difference judgments (Experiment 2), suggesting that the underlying comparison process is the same. We discuss the implications of these findings.

Keywords: Relational Encoding; Object Encoding; Similarity Comparison; Difference Comparison

Introduction

A remarkable feat of human cognition is our ability to represent and reason about relational structure. Relations are an indispensable constituent of both everyday thought and abstract reasoning. We talk about the apple *on* the table (a spatial relation), trace our familial links to a cousin (kinship relations), and think about reciprocity in trade agreements (a relational category; Gentner & Kurtz, 2005). Relations are especially prominent in analogical comparison. According to structure-mapping theory, comparison involves a process of structural alignment based on sharing relational structures (Gentner, 1983; Gentner & Forbus, 2025). By aligning and highlighting the common relations, comparison helps support acquisition of abstract knowledge (Zheng et al., 2025), facilitate solution transfer in problem-solving (Gray & Holyoak, 2021), and lead to scientific discoveries (Dunbar & Blanchette, 2001; Gentner, 2002; Nersessian, 1988). However, decades of research suggest that relational reasoning is more cognitively challenging than that about objects or attributes. The current study adds to this literature, focusing on one source of difficulty: that relations are slower to encode than objects. We also provide novel evidence that this pattern holds for both similarity and difference comparisons.

The difficulty of relational reasoning has long been documented in the developmental literature. For example, Gentner (1988) proposed the relational shift hypothesis: that young learners show an object focus, which gradually shifts to a relational focus as children gain knowledge (Gentner & Rattermann, 1991; Gentner & Boroditsky, 2001; see also Carstensen et al., 2019; Walker et al., 2016). Gentner suggested that the difficult in understanding relational concepts is due in part to their encoding variability. Whereas objects are readily individuated as wholes, relations between objects are often consistent with several different parsings; thus, children must learn the contextually relevant relations (Gentner, 1982; 2005). This knowledge-based shift is found in adults as well. For example, domain experts are more likely to recognize deeper relational patterns than novices (Chi et al., 1988; Rottman et al., 2012).

Converging evidence also comes from studies showing that relational processing is more resource-intensive than object processing. Sophisticated relational reasoning involves binding objects to roles, maintaining and integrating multiple relations into a coherent structure [1], and inhibiting interferences from competing matches. These subprocesses place substantial demands on working memory and inhibitory control (Halford et al., 1998; Hummel & Holyoak, 2005; Simms et al., 2018).

Medin, Goldstone, and Gentner (1990) sought to distinguish relations from objects in a simple perceptual comparison task. In a match-to-sample task, half the participants were asked which one of the two options was more similar to the top standard; the other half were asked which was more different. One option was a partial object match (e.g., sharing a striped circle; Figure 1a, option A), and the other option shared a common relation (e.g., having a uniform pattern) but did not share any objects (Figure 1a, option B). The results revealed a paradoxical asymmetry: people chose the relational match when choosing 'more similar' (as the researchers had predicted) but also when

[1] Neuroimaging results suggest that these subprocesses are handled across a distributed network in the left-lateralized prefrontal areas, which are preferentially recruited during relational processing (Bunge et al., 2009; Green et al., 2010; Holyoak & Monti, 2021; Krawczyk, 2012).

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

choosing “more different”. Medin et al. had no explanation for this same-different asymmetry, though they noted that it is evidence of a clear psychological distinction between relational and object representations.

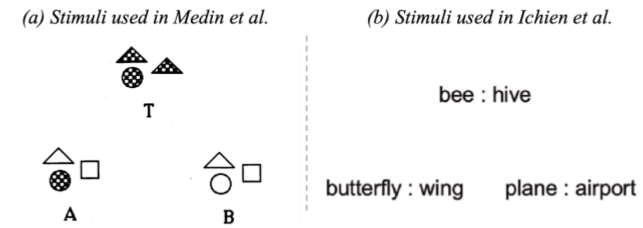

Figure 1: Sample stimuli of Medin et al. (1990) and Ichien et al. (2024) to show an asymmetry between similarity and difference comparisons.

A recent study by Ichien et al. (2024) replicated the original findings of Medin et al. using both perceptual stimuli and verbal stimuli (Figure 1b). Ichien et al. also proposed an account for how the asymmetry arises. Specifically, they argued that (1) processing relations is more cognitively demanding than processing objects or object features, and (2) that processing difference is more complex than processing similarity. These claims are based on prior evidence that ‘*different'* appears to be mentally represented as ‘*not same*’ (in both young children and adults; Hochmann et al., 2016, 2018). This kind of negation adds processing cost (Clark, 1974; Clark & Chase, 1972), and together with the higher demand of processing relations, the result is that comparisons assessing difference are more cognitively complex than comparisons assessing similarity. Ichien et al. further claimed that this asymmetry cannot be accounted for by structure-mapping theory: “Under this alignment hypothesis… assessments of similarity and difference both depend on analogical comparison and involve an identical process of structural alignment… [which] predicts symmetric responding” (p. 14).

We propose an alternative account, which explains the asymmetry by differences in encoding without invoking distinct comparison processes. Our position is that the asymmetry between similarity and difference arises at the encoding stage. We further assert that the same structural-alignment process underlies comparison for both similarity and difference, consistent with structure-mapping theory.

We agree with Ichien et al. (and Hochmann et al., 2016) that *different* is interpreted as *not-same*. This means that the instruction “choose the different one” imposes a greater working-memory load than “choose the similar one”. Thus fewer resources are left for encoding. We next assume that relations are slower to encode than objects and require more attentional resources, as suggested by prior findings (Lovett et al., 2009; Goldstone & Medin, 1994; Zheng et al., 2022).[2] Thus the diminished resources imposed by a difference task will impair relational encoding to a greater degree than object encoding—so a difference task will result in impaired representation of relations (but not objects). In contrast, instructions to choose the most similar do not involve an implicit negative, leaving more resources for encoding the stimulus. This means that when asked “which one is more similar”, people are likely to have encoded both objects and relations. Preferring the relational match, they choose it as most similar. But when asked “which one is more different,” people’s object-dominated representations lead them to respond based mainly on object similarity—so they choose the item with fewer object matches (the relational match) as most different.

This encoding account depends crucially on the claim that relational encoding is slower than object encoding. Here we present two experiments that test this claim. We use matching tasks with a deadline procedure (Simms, 2013; Sloutsky & Yarlas, unpublished manuscript). People are asked to identify a match (either of objects or of relations) from a nonmatch. Critically, we manipulate how long people are allowed to encode the standard. The encoding times span from a brief display of 10ms to a sufficient length of 500ms. We have four predictions.

1. Objects are fast to encode. So performance in the match task will be relatively insensitive to varying encoding times.
2. Encoding relations takes longer. So performance will degrade as encoding time becomes shorter.
3. These patterns will hold for both similarity (Experiment 1) and for difference (Experiment 2) judgments, suggesting the same underlying process.
4. However, performance will be higher overall in the similarity task (Experiment 1) than in the difference task (Experiment 2) because of the additional processing cost imposed by difference judgment. This should be manifested as higher accuracy and/or faster response time.

Experiment 1

Methods

Participant 100 undergraduate students were recruited and were evenly placed into the five conditions (20 each). Participants were recruited through an introductory psychology course at a university in the United States and were compensated with course credits.

Match-to-Sample Task Two matching tasks were used: *relational* (RM) and *object* (OM). On each trial, participants were shown a standard for a fixed amount of time. On the next screen, participants were asked to identify which one of the two alternatives matched the standard. All pictures depict triplets of simple geometric shapes (Figure 2). The OM task involves matching based on the component shapes

[2] To some degree, this follows from the fact that relations (by definition) must apply to their arguments. One cannot encode “*SAME-PATTERN* (A, B)” without having first encoded the objects A and B.

(regardless of the order of the shapes), and the RM task involves matching based on the relational patterns (regardless of the component shapes). For example, for a standard of *square-cross-square*, the correct match would be *cross-square-square* in OM (because of sharing the common objects) and be *triangle-diamond-triangle* in RM (because of sharing the ABA pattern). Importantly, the other option is always a nonmatch (i.e., consisting of neither the same shapes nor the same pattern), so each trial has an unambiguous correct answer.

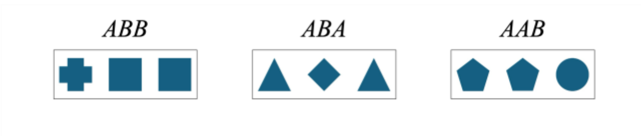

Figure 2: Sample triplets used in Experiments 1 and 2. Triplets were generated from 3 pairs of geometric shapes: *square/cross*, *triangle/diamond*, and *circle/pentagon*, and formed 3 relational patterns: *ABB*, *ABA*, and *AAB*.

The key manipulation was the time allowed for encoding the standard—that is, how long the standard was presented on the screen. Five presentation durations were used based on the results of a pilot study: 10, 30, 50, 100, and 500ms. Figure 3 depicts the procedure for the two matching tasks.

Procedure Participants completed the experiment in a quiet room on the software platform of PsychoPy3 (v2024.2.4; Peirce et al., 2022). The experiment consisted of three phases: demonstration and two blocks of the matching task. One block involved OM and the other involved RM, with order randomized. In all cases, participants were asked to find the match that was more similar (either in relations or in objects) to the standard. Presentation durations were manipulated between-subjects, such that the trials in both blocks used the same presentation duration. Within each block, participants received two practice sets before the test trials.

Demonstration The goal of this phase was to demonstrate the relational and object matches and to practice with key pressing. In two trials (one for each matching task), participants were told, "Sometimes two pictures can be similar if they share the same elements [pattern]," with an example triad of animal triplets. Participants were then told to press 'w' if the left option matched the standard, 'p' if the right option matched the standard, and 'spacebar' for self-paced advancing.

The following procedures were repeated for each of the two blocks: *Practice 1, Practice 2,* and *Test.*

Practice 1 There were 6 practice trials, with order randomized. All trials contained novel shape triplets differing from those used in the test. The format was the same as that for the test (Figure 3), except that the standard was always presented for 1000ms and that feedback was provided following each response. 100% accuracy was required to advance to the next section, or this section would repeat. Everyone reached 100% accuracy within 4 rounds.

Practice 2 The next practice set aimed to prepare participants for the faster encoding time. A new set of 6 practice trials was used, with order randomized. The format was exactly the same as that for the test: the standard was presented for a short duration corresponding to the respective condition. No feedback was provided. This set was only practiced once.

Test There were 18 test trials (3 patterns x 3 pairs of shapes x 2 [which shape was *A* and which was *B*]; Figure 2), with order randomized. The task format was shown in Figure 3. Each trial began with a 'Start' screen, followed by a fixation screen for 500ms. Then, a standard triplet was presented for a fixed amount of duration corresponding to the respective condition (10, 30, 50, 100, and 500ms). The two alternatives—one match and one nonmatch—appeared in the next screen, and participants were asked to identify the match. No feedback was provided. Following each response, participants automatically advanced to the next trial.

Three catch trials were added to the end of each test section with no special prompt. Catch trials consisted of novel shapes used in the two practices and were presented in the test format. They involved an identity match (matching both shapes and pattern) and a nonmatch. The exclusion criterion was predetermined to be more than one error. Only one block from one participant was excluded in this way (an RM block under 10ms condition).

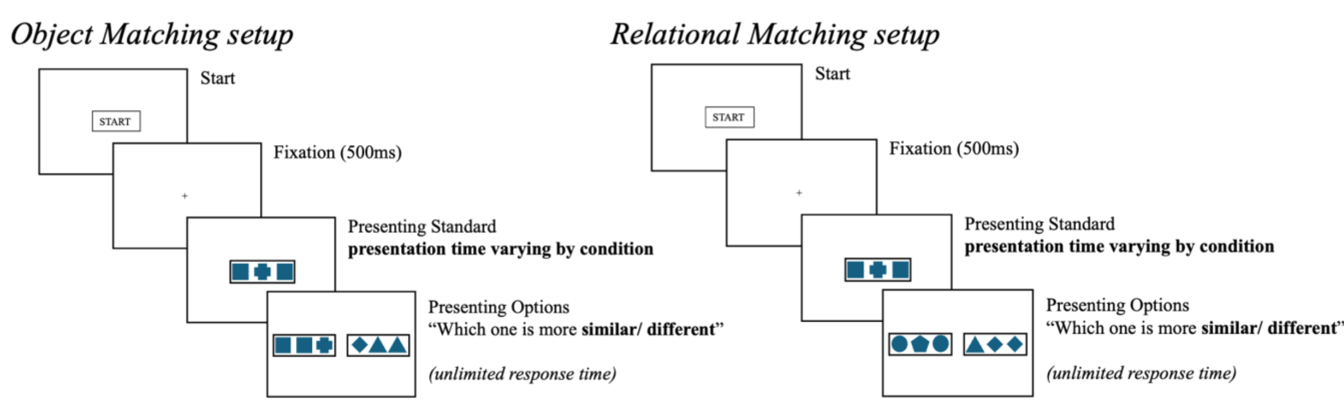

Figure 3: Procedure illustration of Experiments 1 (*Similarity*) and 2 (*Difference*).

Results

The main question of interest was how accuracy for RM versus OM varied across a range of encoding times[3]. To preview, the results bore out our two predictions: (1) the accuracy for OM was relatively insensitive to the varying presentation durations, but (2) the accuracy for RM decreased with shorter durations. We also plotted the response time for correct trials to explore potential speed-accuracy tradeoff.

Overall Trend A power curve was fit to examine the effect of varying encoding time. To do so, we first calculated the average accuracy score by each subject and matching task and log-transformed the scores[4]. The power curve was modeled by using the log-transformed data as the dependent variable of a linear regression. Three independent variables were included: matching task (binary: Relation or Object, with Object being the reference level), presentation duration (ordinal: from 10ms to 500ms), and their interaction.

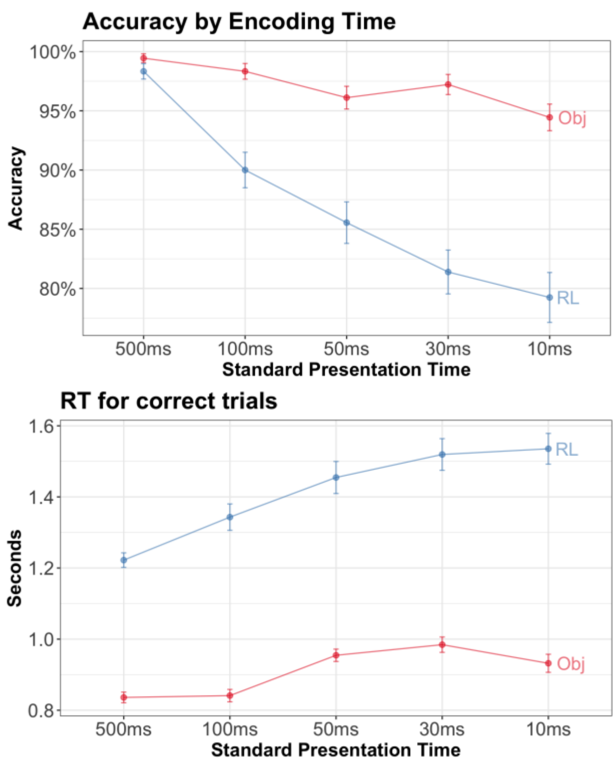

Figure 4: E1 results by presentation duration and matching task: Accuracy (top) and Response time (bottom). Note that the x-axis is from 500ms to 10ms. (RL and Obj stand for Relation and Object, respectively. Error bar represents within-subject standard error)

There was a nonsignificant effect of presentation duration, b = .068, p = .32. Since OM is the reference level, this suggests that for OM we did not see an effect of encoding time, consistent with our first prediction. The main effect of matching type was significant, b = -.17, p < .001, and the negative coefficient suggests that the performance was overall lower in RM than in OM. Importantly, the interaction between matching task and presentation duration was also significant, b = .33, p < .001. This is consistent with our second prediction, that the RM performance would significantly drop as the presentation duration diminishes (Figure 4, top).

Pairwise Comparison Next we compared whether accuracy differed between pairs of presentation durations. Mixed-effect logistic regressions were used to analyze the trial-level accuracy data (1 for correct and 0 otherwise), with two of the durations as the independent variable (a binary factor) and participant as the random effect variable. A significant regression coefficient indicates that the two duration conditions differ in their average accuracy.

Table 1 summarizes the p-values of these regression coefficients. Accuracy was highest at 500ms for both OM and RM. However, while the performance for the rest conditions did not differ much for OM, the pattern for RM showed a rapid decrease, with a possible asymptote toward the shortest durations.

We also used the above model to compare RM performance with OM performance for each duration condition. The results showed that accuracy was lower for RM at every presentation duration except at 500ms (which was at ceiling for both).

Table 1: E1: P-value table for the pairwise comparison of accuracy between conditions.

OM	100ms	50ms	30ms	10ms
500ms	.2	.040*	.036*	.013*
100ms		.2	.32	.054†
50ms			.61	.5
30ms				.22

RM	100ms	50ms	30ms	10ms
500ms	<.001*	<.001*	<.001*	<.001*
100ms		.18	.050*	.001*
50ms			.42	.078†
30ms				.49

Speed-accuracy Tradeoff We explored whether participants sacrificed accuracy for speed in their responses. Figure 4 shows that this was not the case. Participants were in general slower in RM than in OM. Specifically, response times ranged between [0.83, 0.96] seconds for OM and between [1.22, 1.57] seconds for RM. Accuracy ranged between [95.2%, 99.4%] for OM, and between [80.1%, 98.3%] for

[3] We excluded any trial whose response time was three standard deviations longer than the mean response time with respect to the matching task. This constituted 1.56% of all trials in E1 and 1.73% of trials in E2.

[4] Using the original scale did not change the regression patterns in either E1 or E2. In this case, the fitted model was a simple linear regression.

RM (Figure 4). The fact that the performance was both slower and less accurate in RM suggests a robust difficulty for relational encoding.

Experiment 2

E2 involved the same two matching tasks as in E1 but asked for difference judgment instead of similarity judgment. The predictions were the same as for Experiment 1.

Methods

Participant A power simulation with the *paramtest* package (Hughes, 2025) showed that the design of E1 yielded an estimated power of 89% for the interaction effect between matching task and presentation duration. Therefore, 100 undergraduate students were recruited and were evenly placed into the five conditions (20 each). Participants were recruited using the same method as in E1.

Procedure The design was changed for one critical aspect: Participants were asked to look for difference instead of similarity. Specifically, in the matching tasks, participants needed to identify the nonmatch instead of the matching option. The instruction became, “You will see one picture and decide which of the following two is more different from that one,” and “One way two pictures can be different is if they *do not* share the same elements [pattern].” Participants were told to press ‘w’ if the left option was more different and ‘p’ if the right option was more different. All other procedure was the same as in E1.

Results

Data from one block (RM under 30ms) were excluded due to failing the catch trials. To preview, the results paralleled those in E1 and again bore out our two predictions. As the presentation duration diminished, (1) the OM accuracy was largely unaffected, but (2) the RM accuracy decreased steeply. The fact that E1 and E2 showed a parallel pattern provides support for our prediction (3): that similarity comparison and difference comparison share the same underlying process.

Overall Trend The same regression approach was used as in E1. The model showed a nonsignificant effect of presentation duration, b = .043, p = .51, suggesting that varying presentation duration had no effect on accuracy for OM (the reference level). The main effect of matching type was significant and negative, b = -.13, p < .001, which was further qualified by a significant interaction, b = .25, p = .008. These results confirm that the RM accuracy was overall lower than that for OM and decreased as a function of the presentation durations (Figure 5, top).

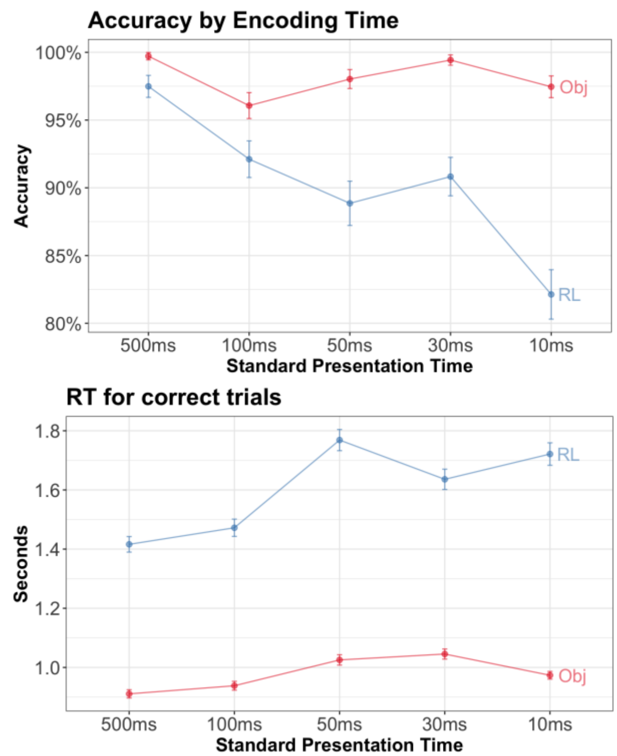

Figure 5: E2 Results by presentation duration and matching task: Accuracy (top) and Response time (bottom). Note that the x-axis is from 500ms to 10ms.

Pairwise Comparison Mixed-effect logistic regressions were used to compare the accuracy between pairs of presentation durations (Table 2). As was the case in E1, accuracy was highest at 500ms for both OM and RM. For shorter durations, while accuracy did not change much for OM, it decreased sharply for RM[5]. Furthermore, comparing the accuracy level between OM and RM revealed lower accuracy for RM at every presentation duration (including 500ms).

Table 2: E2: P-value table for the pairwise comparison of accuracy between conditions.

OM	100ms	50ms	30ms	10ms
500ms	.07†	.096†	.59	.056†
100ms		.62	.25	.75
50ms			.4	.83
30ms				.28

RM	100ms	50ms	30ms	10ms
500ms	.006*	<.001*	.013*	<.001*
100ms		.13	.93	.009*
50ms			.24	.12
30ms				.023*

[5] In Figure 5 top, there seem to be some local nonlinearities in the accuracy patterns for both OM and RM. This is not predicted, but we note that this may be in part driven by individual variations in their ability to detect matches under time constraints. However, as shown in Table 2, these nonlinearities did not reach significance.

Discussion

Two simple perceptual comparison tasks examined people's encoding of relations and objects. The results support our predictions that (1) objects are fast to encode, but (2) encoding relations is slower and takes more resources. Importantly, we saw a close resemblance in patterns for similarity and difference judgments. This supports the ideas that (3) the same process underlies similarity and difference comparisons, though (4) difference judgment takes longer overall, consistent with the idea that there are extra costs in interpreting 'which one is more different.'

The immediate issue driving this research was how to explain the same-different asymmetry found by Medin et al. (1990)—namely, that in a triad similarity task, the relational match is judged to be both more similar and more different than an object match. Ichien et al. (2024) explained this, first, by assuming that interpreting 'which is more different' requires more resources than interpreting 'which is more similar', because *different* is mentally represented as *not-same* (Hochmann et al., 2016). We agree with this assumption. Ichien et al. further assumed that that computing 'different' requires a different and more complex process than computing "similar.' Here we disagree.

It is important to differentiate effects on encoding from effects on processing. We suggest that the same-different asymmetry can be parsimoniously explained by focusing on encoding: (1) Being asked to choose the most different (instead of the most similar) imposes a processing load during encoding of the standard; (2) the reduction in encoding resources is more deleterious for relation than for object, because encoding relations requires more time and resources. These assumptions are borne out in the present results.

In fact, Ichien et al.'s own evidence suggests that the crux of the asymmetry likely resides in encoding. In their Experiment 2, the authors presented people with triads of stories. In both conditions, the standard was presented for a minimum of 10 seconds. In one condition, the two options were then presented simultaneously. Here people showed the typical same-different asymmetry effect. In a second condition, the options were revealed sequentially, and participants were required to read each one for at least 10 seconds. In this condition, with sufficient encoding time for all the materials, the asymmetry between similarity and difference judgments disappeared—evidence that the same-different asymmetry can be explained by differences in encoding time and/or resources, without invoking different comparison processes.

Ichien et al. further argued that because structure-mapping assumes the same process for similarity and difference, it cannot capture the above asymmetry without invoking alignable differences[6]. Again, we disagree. Our contention that the same structural alignment process is used to compute difference as well as similarity is supported by the parallel patterns found between our two experiments, as well as by prior evidence. For example, when asked to name a difference between two things, people are faster for high-similarity pairs than for low-similarity pairs (Gentner & Gunn, 2001; Gentner & Markman, 1994). More tellingly, people are faster to name a difference between two concepts if they have first judged the similarity between the pair—suggesting that the same process of alignment is used for both (Gentner & Gunn, 2001).

In sum, we contend that the same-different asymmetry stems from the encoding process. As laid out in the Introduction, our account is as follows:

- *Difference* is mentally represented as *not-same*. So the instruction for difference ("choose the different one") imposes an extra working memory load over that for similarity. This leaves fewer resources for encoding.
- Encoding relations is slower and takes more resources than encoding objects.
- Thus, the drop in available resources created by a difference task will likely make relational encoding poorer than in a similarity judgment. (so that one is less likely to see the relational commonalities).
- Now consider the triad task. In a similarity task, people will prefer the more satisfying relational match. But in a difference task, people fail to perceive the matching relations. Thus they see the object match as most similar and therefore choose the relational match as more different.

The current study adds to the growing literature examining the characteristics of relational encoding. Despite the prevalence of relations in everyday thought, research has found that relations tend to be encoded more variably than objects (Asmuth & Gentner, 2017; King & Gentner, 2022) and that novice learners often miss the relations altogether (Gentner & Rattermann, 1991; Rottman et al., 2012). We contribute to the discussion by using the deadline procedure to document the real-time encoding profile, which shows that relations are encoded more slowly than objects. This encoding difference has implications for everyday learning. As demonstrated by the asymmetry phenomenon, learners under cognitive load may be likely to settle on the fast-computed object similarities and miss the potentially important relations.

Despite these difficulties, there is evidence that the challenge of relational encoding can be mitigated by further learning. Knowledge accretion and the acquisition of relational language are ways to augment our relational capacity and to facilitate the encoding and processing of relational structure (for reviews, see Gentner, 2010; Goldwater & Schalk, 2016; Richland & Simms, 2015). With the help of these tools, relational encoding can become habitual or even automatic in the case of expertise, supporting the impressive fluency of everyday relational reasoning.

[6] While the distinction between alignable and nonalignable differences is psychologically important (Markman & Gentner, 1993; Sagi et al., 2012; Zheng et al., 2025), we do not think it participates in similarity and difference judgments in the way Ichien et al. described.

References

Asmuth, J., & Gentner, D. (2017). Relational categories are more mutable than entity categories. *The Quarterly Journal of Experimental Psychology, 70*(10), 2007–2025. doi:10.1080/17470218.2016.1219752

Bunge, S. A., Helskog, E. H., & Wendelken, C. (2009). Left, but not right, rostrolateral prefrontal cortex meets a stringent test of the relational integration hypothesis. *Neuroimage*, *46*(1), 338-342.

Carstensen, A., Zhang, J., Heyman, G. D., Fu, G., Lee, K., & Walker, C. M. (2019). Context shapes early diversity in abstract thought. *Proceedings of the National Academy of Sciences, 116*(28), 13891-13896.

Chi, M.T.H., Glaser, R., & Farr, M.J. (Eds.). (1988). The Nature of Expertise (1st ed.). Psychology Press.

Clark, H. H. (1974). The chronometric study of meaning components. Problems actuels en psycholinguistique (pp. 489–505). *Paris: Editions du Centre National de la Recherche Scientifique*.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive psychology, 3*(3), 472-517.

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: *The case of analogy. Trends in cognitive sciences, 5*(8), 334-339.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *BBN report; no. 4854*.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science, 7*(2), 155-170.

Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp.21-39). New York: Kluwer Academic/ Plenum Publisher.

Gentner, D. (2005). The development of relational category knowledge. In *Building object categories in developmental time* (pp. 263-294). Psychology Press.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive science, 34*(5), 752-775.

Gentner, D., & Forbus, K. (2025). The Structure-Mapping Engine: A Multidecade Interaction Between Psychology and Artificial Intelligence. *Current Directions in Psychological Science*, 09637214251395678.

Gentner, D., & Gunn, V. (2001). Structural alignment facilitates the noticing of differences. *Memory & cognition, 29*(4), 565-577.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W.-k. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin (pp. 151–175). *American Psychological Association*.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological science, 5*(3), 152-158.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215-256). Cambridge, UK: Cambridge University Press.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. *Perspectives on language and thought: Interrelations in development*, 225.

Goldstone, R. L., & Medin, D. L. (1994). Time course of comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 29.

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, *142*(7), 729.

Gray, M. E., & Holyoak, K. J. (2021). Teaching by analogy: From theory to practice. *Mind, Brain, and Education, 15*(3), 250-263.

Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral cortex, 20*(1), 70-76.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and brain sciences, 21*(6), 803-831.

Hochmann, J. R., Carey, S., & Mehler, J. (2018). Infants learn a rule predicated on the relation same but fail to simultaneously learn a rule predicated on the relation different. *Cognition, 177*, 49-57.

Hochmann, J. R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match-and non-match-to-sample. *Cognitive psychology, 86*, 87-111.

Holyoak, K. J., & Monti, M. M. (2021). Relational integration in the human brain: A review and synthesis. *Journal of cognitive neuroscience*, *33*(3), 341-356.

Jeff Hughes (2025). paramtest: Run a Function Iteratively While Varying Parameters. R package version 0.1.1. https://cran.r-project.org/package=paramtest

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Psychological Science*, *14*(3), 153-157.

Ichien, N., Lin, N., Holyoak, K. J., & Lu, H. (2024). Cognitive complexity explains processing asymmetry in judgments of similarity versus difference. *Cognitive Psychology, 151*, 101661.

Jamrozik, A., & Gentner, D. (2020). Relational labeling unlocks inert knowledge. *Cognition, 196*, 104146.

King, D., & Gentner, D. (2022). Verb metaphoric extension under semantic strain. *Cognitive Science, 46*(5), e13141.

Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain research*, *1428*, 13-23.

Lovett, A., Gentner, D., Forbus, K., & Sagi, E. (2009). Using analogical mapping to simulate time-course phenomena in perceptual similarity. *Cognitive Systems Research, 10*(3), 216-228.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1990). Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Science, 1*(1), 64-69.

Nersessian, N. J. (1988, January). Reasoning from imagery and analogy in scientific concept formation. In PSA: *Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1988, No. 1, pp. 41-47). Cambridge University Press.

Peirce, J. W., Hirst, R. J. & MacAskill, M. R. (2022). Building Experiments in PsychoPy. 2nd Edn London: Sage.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 177-192.

Rottman, B. M., Gentner, D., & Goldwater, M. B. (2012). Causal systems categories: Differences in novice and expert categorization of causal phenomena. *Cognitive science, 36*(5), 919-932.

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive science, 36*(6), 1019-1050.

Simms, N. K. (2013). Overcoming the object bias in analogical development: Encoding time, allocation of attention, and relational knowledge (Doctoral dissertation, Northwestern University).

Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology, 166*, 160-177.

Sloutsky, V. M., & Yarlas, A. S. (unpublished manuscript). Processing of information structure: Mental representations of elements and relations.

Walker, C. M., Bridgers, S., & Gopnik, A. (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. *Cognition, 156*, 30-40.

Zheng, Y. S., Goldwater, M., & Gentner, D. (2025). Structural Alignment and Linguistic Contrast Help Children Learn a Key Principle of Spatial Construction. *Cognitive Science, 49*(12), e70149.

Zheng, Y., Matlen, B., & Gentner, D. (2022). Spatial alignment facilitates visual comparison in children. *Cognitive Science, 46*(8), e13182.