Title
Comparison and Explanation in Learning Causal System Categories

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Pavlova, Margarita V.
Greenhoot, Andrea Follmer

Publication Date
2023

Peer reviewed

Comparison and Explanation in Learning Causal System Categories

Margarita Pavlova (mvpavlova@nbu.bg)
Department of Cognitive Science and Psychology, New Bulgarian University, 21 Montevideo Street,
Sofia 1618, Bulgaria

Andrea Follmer Greenhoot (dea@ku.edu)
Department of Psychology, University of Kansas, 1450 Jayhawk Blvd,
Lawrence, KS 66045 USA

Abstract

The ability to notice relational patterns across situations is crucial to learning. An important question is how people build transferable and generalizable knowledge by learning from examples. The study investigated the conditions that make comparison and explanation beneficial to learning by comparing learning outcomes of engaging in explanation or comparison in a categorization task and examining how varying degrees of instructional support affect these two processes. The results showed an advantage of comparison over explanation; however, this was specific to combination of relational labels and definitions and prompts to compare. These results add to existing research and extend our understanding of how best to support college students learning relatively difficult material. The findings also inform ways educators can support learning by developing instructional designs that support learning through analogical reasoning.

Keywords: comparison; explanation; learning; transfer

Introduction

Research on analogical comparison (comparing or aligning two or more analogous cases) and explanation (answering why questions) has shown that both processes appear to aid knowledge transfer and are thought to support a capacity to make inferences and generalizations that go beyond the surface-level similarities between a previously learned and novel information (Hummel et al., 2014), but they have different instructional implications. Understanding when and how comparison and explanation foster learning and subsequent transfer and what instructional approaches might support students' acquisition of relational-structural information across domains, is an important step in expanding our understanding of these cognitive processes and it can inform good practices in education.

Although explanation and comparison aid transferable learning, they might rely on different cognitive mechanisms. One way in which comparison is thought to support learning is through difference detection. In addition to highlighting commonalities between two analogs, structural alignment also highlights alignable differences (i.e., which belong to the common relational structure). Since learners typically analyze these differences, comparison can enable them to notice deeper similarities and supports the formation of an abstract relational schema despite surface-level differences between compared cases (Gentner et al., 2009). One way in which explanation is thought to support learning is by increasing metacognitive awareness and providing a general boost to attention and cognitive engagement (Chi, 2000, 2009; Roy & Chi, 2005). Another way, according to the unification and subsumption account, is that explanation leads learners to *selectively* prefer unifying patterns (Legare & Lombrozo, 2014; Lombrozo & Carey, 2006; Walker et al., 2014; Williams & Lombrozo, 2010). Thus, explanation guides people to interpret the examples in terms of broader patterns.

While considerable work shows that comparison and explanation *individually* can support transferable learning, a few studies investigate these processes in the same experimental task. Nokes-Malach et al. (2013) found that college students who explained solutions to worked physics examples outperformed students who compared pairs of examples on near transfer problems, however, both groups performed comparably on far transfer problems and outperformed the control group. Richey et al. (2015) found mixed efficiency of explanation and comparison on college students' learning about electricity. Specifically, students who explained worked examples and those who compared them performed worse than the control group (who read instructional explications) on near transfer, whereas on intermediate transfer, students who explained outperformed those who compared. All groups struggled with the far transfer task, and on a task measuring preparation for future learning, the students in the control group outperformed those who compared. Gadgil et al. (2012) compared the effect of strategies aimed at promoting mental model revision (e.g., comparing one's own mental model to an expert model) and strategies aimed at promoting revising false beliefs (e.g., self-explaining the expert model alone) on college students' knowledge of the human circulatory system. They found that students in the comparison condition were more likely to achieve conceptual change towards the expert model than those in the other conditions. While these studies offer insight into the conditions that make comparison and explanation beneficial to learning, it is still unclear what are the mechanisms by which these two processes foster learning and transfer.

Several studies offer insights about how explanation and comparison might interact to support learning. Kurtz et al. (2001) asked college students to learn from dissimilar examples of heat flow and found that relative to students who

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

only described and explained the scenarios, students who also compared them found more similarities between them. In addition, comparison was most effective when students were asked to list corresponding similarities between the two examples. Sidney et al. (2015) compared college students' learning about fraction division from worked examples; students were assigned in either explanation, comparison (indicating whether there are similarities and differences), both explanation and comparison, or a control condition. The results showed that procedural learning did not differ among conditions; however, participants who explained the examples showed more conceptual knowledge gain than those who compared them; the combined condition lead students to report more similarities and differences; students who only compared cases and were not asked to explain, showed less learning. Taken together, these results indicate that explanation fostered conceptual learning but simply noticing similarities and differences without making sense of them did not increase learning. Goldwater & Gentner (2015) asked participants to learn about causal systems by reading a causal label or a full explication of how the example fit the system; then some participants aligned pairs of examples belonging to the same system. The results showed that participants who compared pairs were more likely to sort new examples to match the causal system. Moreover, comparison was most beneficial when it was combined with the full explication. Edwards et al. (2019) asked participants to categorize robots according to rules by either engaging in explanation or comparison or neither and learning from pairs of category members (pair-wise comparison) or from all members of the category (group-level comparison). Overall, the results showed that prompts to explain increased the likelihood of discovering category rules. In addition, prompts to explain invited participants to compare to an even greater degree than actual prompts to compare did. Further, group-level comparison (but not pair-wise comparison) partially mediated the effect of explanation on rule discovery. These results led the researchers to propose that explanation supports category learning by recruiting comparisons that support the identification of abstract commonalities.

An important issue for education is that spontaneous analogical transfer is difficult, especially when the two domains are dissimilar at the surface level (Chi et al., 1981; Forbus et al., 1995; Gentner et al., 1993; Gick & Holyoak, 1980, 1983). Importantly, transfer depends on how a learner would make sense of a new problem in terms of their prior knowledge (Goldwater & Schalk, 2016). The ability to transfer knowledge to a new case relies on both encoding the initial case, particularly in terms of relational information, as well as retrieving an appropriate base at the time of transfer. Thus, sometimes, failure to transfer might be caused by insufficient learning of the to-be-transferred concepts (Chi & VanLehn, 2012).

One way to aid transfer is to support the encoding of the study material. Evidence from analogical reasoning research shows that providing a label for the relational schema or solution principle facilitates comparison (Christie & Gentner, 2010; Son et al., 2010) and fosters transfer (Christie & Gentner, 2014; Gentner et al., 2011; Gick & Holyoak, 1983; Jamrozik & Gentner, 2020; Kubricht et al., 2017; Loewenstein & Gentner, 2005; Son et al., 2010). However, it is unclear whether providing more context-specific (e.g., full explication of an example) or more abstract information (e.g., definition of a principle) would support or hinder building an abstract schema to transfer to novel situations, as there are conflicting views on the effectiveness of definitions particularly in learning novel complex material. Forbus and Gentner (1986) proposed that the early stages of learning are governed by rich, context-specific representations thereby necessitating the need for learning from concrete examples rather than abstract schemas. Abstract definitions might be more poorly understood than specific cases and call for many different interpretations of a relational abstraction, meaning that people might encode the definition in a way that is incompatible with the example (Gentner et al., 2004) and completely abstract materials interfere with comprehension (Day & Goldstone, 2012). On the other hand, an appropriate abstract definition directly provides the relational schema and can aid schema induction and subsequent transfer (Eiriksdottir & Catrambone, 2011; Gick & Holyoak, 1983; Kubricht et al., 2017). Further, providing abstract labels of subgoals (principles) (Catrambone, 1998) or one's own subgoal labels as scaffolds (Margulieux & Catrambone, 2021) improves later problem-solving performance. Providing a full explication might support building a rich, well-connected representation of the base analog, because it might increase comprehension and provides an opportunity for deep understanding (Chi & VanLehn, 2012). Moreover, prior knowledge interacts with the ability to extract relevant relations. For example, experts are thought to habitually encode new examples in the terms of key relational principles which contributes to their ability to retrieve relational matches (Goldwater et al., 2021). Conversely, contextual information can impede learning and transfer because it might distract from the structural information (Day & Goldstone, 2012; Ross, 1989). Indeed, reducing the "richness" of the learning material facilitates alignment and transfer (Markman & Gentner, 1993). To our knowledge, the effectiveness of different levels of support has not been systematically studied and much remains unknown about the conditions in which various degrees of support can promote or hinder learning and transfer.

In sum, substantial research shows that both explanation and comparison support learning and transfer; however, there still seems to be some inconsistency regarding whether explanation or comparison is more beneficial to learning and the degree to which they recruit one another. The current research investigates when and how comparison and explanation support learning. Additionally, the study aims to investigate the role of instructional support in learning and transfer as well as how different levels of instructional support interact with cognitive strategies to foster abstraction and generalization. We hypothesize that both explanation and comparison as engaging in active and constructive learning

(Fonseca & Chi, 2010) would support transfer of relational-causal information. Regarding the levels of support, we expect to replicate previous findings on the role of relational labels, and do not have specific expectations regarding definitions and explications.

Method

Participants

A total of 320 undergraduate students from the University of Kansas (M_{age} = 19.2, SD_{age} = 1.259; 131 female, 131 male, 2 non-binary, 56 did not respond) participated in exchange of course credit. Data from additional 421 participants were excluded because they failed at least one attention check explicitly asking them to sort 2 examples in the AST under a specific causal system. The study was conducted online via Qualtrics, and data were collected between Summer 2020 and Spring 2022.

Design and Procedure

A 3 (Training: Explanation, Comparison, No Training) X 3 (Level of Support: Label, Definition, Explication) + 1 (Baseline) design was used, thus yielding 10 conditions with all variables manipulated between participants. Participants in all but the Baseline condition were presented with ten short examples depicting five causal systems (common cause, common effect, causal chain, positive feedback, or negative feedback). Each example was accompanied by either a label, a label and a definition, or a label, definition, and explication. After reading the examples, participants were presented with the same examples either sequentially and were asked to explain how each fit a given causal system (Explanation conditions) or in pairs and were asked to describe the key parallels between them (Comparison conditions). Next, all participants proceeded to the assessment materials where they solved a sorting task and self-reported the amount of comparison and explanation during the Learning Phase.

Materials

Instruction Materials The instructional materials consisted of ten short examples depicting causal systems from two domains (e.g., electrical engineering and biodiversity; adapted from Rottman et al., 2012). In the first part of the Learning Phase, participants read the examples and the label/definition/explication that followed each of them. In the second part of the Learning Phase, participants received either an explanation prompt (e.g., "In your own words, explain why this description is an example of a [positive feedback] system.") or a comparison prompt (e.g., "These two are examples of a [positive feedback] system. What are the key parallels between these two phenomena?") and wrote their responses in a box provided on the screen. Next, they proceeded to the Transfer Phase consisting of the assessment materials.

Assessment Materials Assessment materials consisted of the Ambiguous Sorting Task (AST; adapted from Rottman et al., 2012). The AST consisted of 20 example phenomena composing a matrix of five causal systems crossed with five content domains. Because there were two types of sorts (e.g., according to domain or causal system), participants sorted the examples twice.

We also collected amount of self-reported comparison and explanation during the Learning Phase on a 7-point Likert scale. Finally, to examine whether prompts to compare and explain elicited comparison among the training groups, the explanations and comparisons students generated were coded for comparison. For each causal system, we coded whether students generated at least one comparison. Because we were interested in the quality of the comparison, we coded for broad and deep comparisons. Broad comparisons were cases in which a student is referring to a previous example or to both example phenomena. Deep comparisons were cases in which a student is explicitly mapping corresponding elements between two examples.

Results

A quasi-randomized order for the example phenomena in the AST was used, thus having two orders. A series of independent samples t-tests showed that this order predicted systematic differences in the types of sorts such that in the first sorting, participants who received Order 2 (N = 166, M_{causal} = .215, SD = .185), sorted more causally ($t(318)$ = 2.132, p = .034) than those who received Order 1 (N = 154, M_{causal} = .172, SD = .176). Participants who received Order 1 (M_{domain} = .498, SD = .206) sorted more by domain ($t(318)$ = 2.433, p = .016) than those who received Order 2 (M_{domain} = .443, SD = .193). There were no differences between the two orders in the amount of error sorts across the two sorting's, nor in the amount of causal and domain sorts in the second sorting. All subsequent analyses controlled for Order.

Types of Sorts in the AST

We first compared the rate of causal and domain sorts between the first and the second sorting among all groups, and among just the training groups. The second sorting was included to allow participants to show sensitivity to the causal-relational similarity across examples, even when their first sorting might have been dominated by domain sorts. Two separate repeated measures ANOVAs on the rate of causal and domain sorts with *Sorting Set* as within-subjects factor and *Condition* and *Order* as between-subjects factors found a main effect of sorting set such that there were fewer causal sorts in the 2nd sorting than the 1st ($F(1,309)$ = 8.778, p = .003, η_p^2 = .028) as well as more domain sorts ($F(1,309)$ = 9.233, p = .003, η_p^2 = .029). Additionally, the model found a main effect of condition in which students in the Baseline condition sorted less by causal system ($F(9,300)$ = 4.256, p < .001, η^2 = .081) and more by domain ($F(9,309)$ = 4.588, p < .001, η_p^2 = .118) than students who were prompted to compare and received either a label or a definition. Table 1 contains the results from the post-hoc analyses.

Next, we compared the performance among the 9 training groups. A repeated measures ANOVA on the rate of **causal**

sorts, with *Sorting Set* as within-subjects factor and *Support, Training,* and *Order* as between-subjects factors found a main effect of Sorting Set, $F(1,272) = 8.630$, $p = .004$, $\eta_p^2 = .031$; the rate of causal sorting was reliably less frequent in the 2nd sorting ($M = .173$) than in the first ($M = .205$), $t(280) = 2.938$, $p = .004$. There was also a main effect of Training, $F(2,272) = 7.083$, $p = .001$, $\eta_p^2 = .05$. Students who were prompted to compare sorted more by causal system than students who did not receive any training, $t(194) = -3.746$, $p < .001$. There was no main effect of Support, $F(2,272) = 1.340$, $p = .263$, no main effect of Order, $F(1,272) = 3.119$, $p = .078$. There was no significant interaction between Sorting Set and Support, $F(2,272) = 1.024$, $p = .360$, nor between Sorting Set and Training, $F(2,272) = 2.951$, $p = .054$, nor a significant triple interaction between Sorting Set, Support, and Training, $F(4,272) = .856$, $p = .491$.

Another repeated measures ANOVA on the rate of **domain sorts** between the *Sorting Set* as within-subjects factor and S*upport, Training,* and *Order* as between-subjects factors found a main effect of Sorting Set, $F(1,272) = 10.568$, $p = .001$, $\eta_p^2 = .037$. The rate of domain sorting was reliably more frequent in the second sorting ($M = .507$) than in the first ($M = .455$), $t(280) = -3.251$, $p = .001$). There was also a main effect of Training, $F(2,272) = 13.476$, $p < .001$, $\eta_p^2 = .09$. Students who were prompted to compare sorted less by domain than students who did not receive any training, $t(194) = 25.171$, $p < .001$ and whose who were prompted to explain, $t(182) = 2.117$, $p = .035$. The students who were prompted to explain also sorted less by domain that the students who did not receive any training, $t(183) = 2.906$, $p = .008$. There was also a main effect of Order, $F(1, 272) = 5.097$, $p = .025$, $\eta_p^2 = .018$. Students who received Order 1 sorted more by domain that students who received Order 2, $t(280) = 2.258$, $p = .025$. There was no main effect of Support, $F(2,272) = 1.99$, $p = .139$. There was no significant interaction between Sorting Set and Support, $F(2,272) = 2.321$, $p = .1$, nor between Sorting Set and Training, $F(2,272) = 2.068$, $p = .128$, nor a significant triple interaction between Sorting Set, Support, and Training, $F(4,272) = .1.041$, $p = .387$.

Self-reported Comparisons and Explanations

The analyses of the sorts in AST indicated that students who were prompted to compare sorted more causally than students who did not receive training and those who were prompted to explain. However, as reviewed at the beginning of this paper, comparison and explanation sometimes recruit each other. Thus, an interesting question is whether prompts to compare would elicit explanation processing and whether prompts to explain would elicit comparison processing. Relatedly, we sought to examine whether these prompts elicit comparisons and/or explanations. To determine what processes the prompts stimulated, we analyzed students' self-reports of the frequency with which they engaged in comparison and explanation during the learning phase.

A 3 *(Support: Label, Definition, Explication)* X 3 *(Training: No Training, Compare, Explain)* ANOVA on the frequency of **self-reported comparison** revealed no significant main effects or interaction (support: $F(2,230) = .826$, $p = .439$, $\eta^2 = .007$; training: $F(2,230) = .574$, $p = .564$, $\eta^2 = .005$; interaction: $F(4,230) = .887$, $p = .472$, $\eta^2 = .015$). Another 3 *(Support: Label, Definition, Explication)* X 3 *(Training: No Training, Compare, Explain)* ANOVA on the amount of **self-reported explanation** also revealed no main effects or interaction (support: $F(2,230) = .139$, $p = .871$, $\eta^2 = .001$; training: $F(2,230) = 1.082$, $p = .341$, $\eta^2 = .009$; interaction: $F(4,230) = .530$, $p = .714$, $\eta^2 = .009$). Generally, students reported relatively high rates of both comparison and explanation.

Rate of Comparison During Learning

Three groups received only support but did not receive further training, while the remaining six groups received prompts to compare or explain. For these six groups, we collected verbal protocols of the explanations and comparisons students generated during learning and looked at the rate of broad and deep comparisons. There were 184 participants: 87 were prompted to explain and 97 were prompted to compare; 58 received a label, 63 received a definition, and 63 received an explication. A total of 92 students generated broad comparisons, and 34 of them generated deep comparisons.

A 3 *(Support: Label, Definition, or Explication)* X 2 *(Training: Compare or Explain)* ANOVA on the amount of **broad comparisons** found a main effect of Training, $F(1,178) = 162.725$, $p < .001$, $\eta_p^2 = .478$. Students who were prompted to compare generated significantly more broad comparisons than students who were prompted to explain, $t(182) = -12.756$, $p < .001$. There was no main effect of Support, $F(2,178) = 2.018$, $p = .136$. There was a significant interaction between Support and Training, $F(2,178) = 3.798$, $p = .024$, $\eta_p^2 = .041$, such that students in all comparison groups outperformed students in all explanation groups. In addition, students who were prompted to compare and received explications generated more broad comparisons than students who were prompted to compare and received definitions, $t(65) = -3.324$, $p = .016$.

Another 3 *(Support: Label, Definition, or Explication)* X 2 *(Training: Compare or Explain)* ANOVA on the amount of **deep comparisons** found a main effect of Training, $F(1,178) = 31.220$, $p < .001$, $\eta_p^2 = .149$. Students who were prompted to compare generated significantly more deep comparisons than students who were prompted to explain, $t(182) = -12.756$, $p < .001$. In fact, none of the students in the explanation groups generated any deep comparisons. There was no main effect of Support, $F(2,178) = .920$, $p = .4$, and no interaction between Support and Training, $F(2,178) = .920$, $p = .4$.

Encouraged by this, we next asked whether individual differences in generating comparisons during learning predicts performance on the transfer tasks across all six groups. However, generating broad or deep comparisons did not correlate significantly with sorting causally in the AST (broad: $r = .007$, $p = .929$; deep: $r = -.039$, $p = .528$).

Table 1: Post-hoc comparisons of Baseline versus the 9 training groups, Panel a) Causal Sorts; Panel b) Domain Sorts

a)

	Causal Sorts			
	Mean Difference	**SE**	**t**	***Cohen's d***
L & C	-0.150	0.034	-4.379***	-0.910
L & E	-0.088	0.036	-2.468	-0.533
L only	-0.031	0.035	-0.894	-0.188
D & C	-0.165	0.033	-4.924***	-0.998
D & E	-0.085	0.035	-2.435	-0.515
D only	-0.047	0.033	-1.414	-0.283
FE & C	-0.066	0.034	-1.946	-0.401
FE & E	-0.066	0.034	-1.912	-0.397
FE only	-0.062	0.034	-1.827	-0.376

b)

	Domain Sorts			
	Mean Difference	**SE**	**t**	***Cohen's d***
L & C	0.133	0.039	3.435*	0.650
L & E	0.081	0.040	2.005	0.395
L only	0.002	0.039	0.057	0.011
D & C	0.180	0.038	4.785***	0.883
D & E	0.115	0.039	2.908	0.561
D only	0.053	0.037	1.437	0.262
FE & C	0.125	0.038	3.251	0.610
FE & E	0.096	0.039	2.475	0.468
FE only	0.034	0.038	0.888	0.167

* $p < .05$, ** $p < .01$, *** $p < .001$

Note. P-value corrected using the Bonferroni method. L = label, D = Definition, FE = Full Explication, C = Compare, E = Explain

Discussion

The study examined the effects of engaging in comparison and explanation, along with the level of support provided for learning the to-be-transferred material on learning and transfer of high-level causal system categories, such as positive feedback loop. Prior work suggests that these categories are not salient to college students without advanced training (Rottman et al., 2012). Our goal was to investigate the sorts of learning processes that facilitate students' ability to recognize these patterns in the world, and to explore how different levels of instructional support might foster or hinder these processes.

We hypothesized that explanation and comparison, as engaging in active and constructive learning (Fonseca & Chi, 2010), would support transfer of relational-causal information. However, the overall results revealed learning advantages only for the Comparison groups. Compared to the baseline, students who were prompted to compare and received either causal labels or labels and definitions with the learning examples sorted more causally and less by content domain. Similarly, we found significant effects of training: students who were prompted to compare sorted more causally than students who did not receive training or who were prompted to explain. Taken together, these findings suggest that prompts to compare were more effective than prompts to explain in fostering students' attention to the relational structure in the learning examples. Additional evidence for this interpretation comes from the finding that students who were prompted to compare sorted less by content domain than students who were prompted to explain and those who did not receive training.

There were no differences in the frequency of self-reported comparisons and explanations between the training groups: students reported relatively high levels of explanation and comparison and seemed to have engaged equally in both processes. However, examining the actual frequency of comparisons generated in students' written responses during learning showed that students in the comparison groups produced more comparisons than students in the explanation groups, and only students in the comparison groups generated deep comparisons (i.e., that include structural alignment). However, these individual differences in generated comparisons did not predict transfer performance. In fact, comparisons were not frequent: 79 students in the comparison groups, and 13 in the explanation groups generated broad comparisons, and 34 in the comparison groups generated deep comparisons. It is possible that while engaging with the examples facilitated students' ability to notice and mention the similarity between them, it did not provide enough support for spontaneous structural alignment. We expected the deep comparisons to correlate with performance and find the lack of a correlation surprising.

For broad comparisons, there was an interaction between support and training: students who were prompted to compare and received explications produced more broad comparisons. It is possible that receiving the explication improved students' comprehension of the examples themselves, thus allowing them to be reminded of other analogous examples. However, the fact that we did not see improved transfer particularly in the comparison and explication group raises a question of how transferable that understanding was. It is possible that while the explication increased comprehension of the examples, it did so in a way that did not promote the abstraction of the causal schema and thus hindered subsequent transfer. Conversely, receiving labels or definitions might have promoted more abstract learning that while did not promote more broad comparisons, generally, led to improved transfer.

We did not find consistent evidence of the effect of differing levels of instructional support. In fact, only labels and definitions, when combined with comparison, lead to more casual and fewer domain sorts. This result adds to existing evidence that relational labels invite comparison and improve learning and transfer (Christie & Gentner, 2010; Gentner et al., 2009; Jamrozik & Gentner, 2020; Kotovsky & Gentner, 1996; Novick, 1988; Son et al., 2010). Bowdle & Gentner (2005) differentiate between horizontal alignment (aligning representations that are at the same level of abstraction, e.g., two examples) and vertical alignment (aligning representation that are at different levels of abstraction, e.g., an example and a label). Here, the representation of the meaning of a label is more abstract than the representation of a full explication. Arguably, the representation of a definition would also be more abstract. Therefore, aligning an example with a label (or a definition) is computationally less expensive than aligning an example and the full explication, because there are fewer extraneous (i.e., surface-level) matches to handle. Moreover, extraneous information can impede learning and transfer because the concrete details of the example may interfere with the students' ability to transfer the causal-relational structure to novel situations even if these details promote comprehension of the example. Prior work suggests that providing simple rather than rich training materials can facilitate mapping and transfer (Markman & Gentner, 1993; Rattermann & Gentner, 1998; Son & Goldstone, 2009). Furthermore, labels and definitions provide consistent language across examples which increases the likelihood that two things that are named with the same label or short definition would be encoded as members of the category and will be retrieved at transfer. Relatedly, Snoddy and Kurtz (2017) found that a study manipulation thought to foster learning via category-building promoted spontaneous analogical transfer. Considering their "category status hypothesis", it might be possible that students who received labels or definitions with the learning examples abstracted the relational category they belong to and thus transferred it to novel examples. However, a more systematic investigation of encoding is necessary to further shed light on the role of encoding in learning and transfer. Regarding the results that students who received more support did not perform better, it is possible that, receiving explications has preempted their own active processing of the learning materials (Gerjets et al., 2006); see Marguleux & Catrambone (2019) for related evidence of optimal combination of support in learning subgoals. Additionally, the presence of full explications might have obviated the need for participants to generate explanations of their own, thus diminishing their learning (Bertsch et al., 2007; Freeman et al., 2014; Theobald et al., 2020).

The current findings are compatible with prior work demonstrating that engaging in comparison supports learning and fosters transfer (Gadgil et al., 2012; Gick & Holyoak, 1980, 1983; Goldwater & Gentner, 2015; Kurtz et al., 2001; Loewenstein et al., 2003). While other work shows that engaging in explanation is beneficial to learning (Chi, 2000; Lombrozo, 2012; Rittle-Johnson & Loehr, 2017), sometimes even more beneficial than comparison (Nokes-Malach et al., 2013; Richey et al., 2015; Sidney et al., 2015), the current results do not seem to support that. It is possible that the students in our study engaged in suboptimal explanations. The benefit of explanation found in prior research generally has been attributed to the quality of explanations generated by participants (Brown & Kane, 1988; Chi et al., 1989; Cho & Jonassen, 2012; Crowley & Siegler, 1999; de Koning et al., 2010; Rittle-Johnson et al., 2015). While we did not investigate the quality of the student generated explanations, it is possible that students who generated good quality explanations would show greater transfer, and future work could explore this possibility. Furthermore, other research suggests that one mechanism via which explanation may support learning is by inviting comparison (e.g., Chin-Parker & Bradner, 2010, 2017; Edwards et al., 2019; Hummel et al., 2014; Needham & Begg, 1991). It is possible that students who were prompted to explain focused on context-specific or idiosyncratic information and thus did not attend to the relational structure in a way that could allow them to recognize it in analogous cases.

Our findings contrast with those of Goldwater and Gentner (2015), who examined the effects of comparison instructions combined with either labels or full explications and found that the combination of explication and structural alignment lead to the highest rate of causal sorting. One reason for the different results obtained in the current study may be that we did not scaffold the alignment the way Goldwater and Gentner (2015) did. The current research used a more general prompt inviting participants to focus on the key parallels between the cases but did not require them to find *corresponding* elements. It is possible that a more scaffolded procedure would increase the amount and quality of the comparison.

The current findings add to existing research showing that comparison can lead to relational understanding that facilitates transfer. Furthermore, the results align with other work arguing for a broader role of analogical reasoning in learning and transfer (Day & Goldstone, 2012; Doumas & Hummel, 2013; Gentner, 2003, 2010; Goldwater & Schalk, 2016). The AST designed by Rottman et al. (2012) tests transfer as a more generalized ability to notice key patterns. Thus, the current findings add to existing research that the effects of analogical reasoning go beyond inference projection (Gentner, 2010; Goldwater & Gentner, 2015; Kurtz et al., 2001) and are consistent with the argument that the mutual structural alignment of examples supports deriving causal system abstractions that aid subsequent transfer by supporting the uniform encoding of relations (Gentner, 2010).

The current findings have implications for educational practice and pedagogical strategies that can promote transferable learning and adds to a body of work suggesting that supporting analogical comparison could improve learning in educational settings, particularly related to math and science (Alfieri et al., 2013; Gentner et al., 2016; Richland et al., 2007; Rittle-Johnson & Star, 2009; Thompson & Opfer, 2010). Furthermore, the findings related to instructional support have implications for understanding the appropriate level of detail that should accompany the cases being learned, at least for college students.

The current study has some methodological limitations that need consideration. Specifically, conducting the study online might have affected the students' engagement which could have attenuated learning effects of our manipulation. Future studies should have better control to ensure that participants engage with the study materials. Additionally, we plan to systematically study the role of differing levels of instructional support in encoding the learning materials and achieving transferable learning. Last, we hope to explore the benefits and drawbacks of learning from within- versus cross-domain examples more systematically in future work.

Acknowledgements

We are grateful to the members of the MAD lab at the University of Kansas for their help with data collection and to Art Markman, Micah Goldwater, Penka Hristova, and Robyn Kelton for discussions regarding the study design and earlier versions of this text. We also thank the anonymous reviewers for their valuable comments and suggestions.

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.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, *35*(2), 201–210.

Bowdle, B. F., & Gentner, D. (2005). The Career of Metaphor. *Psychological Review*, *112*(1), 193–216.

Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology*, *20*(4), 493–523.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, *127*(4), 355–376.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology*. Lawrence Erlbaum Associates.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science*, *13*(2), 145–182.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188.

Chin-Parker, S., & Bradner, A. (2010). Background shifts affect explanatory style: How a pragmatic theory of explanation accounts for background effects in the generation of explanations. *Cognitive Processing*, *11*(3), 227–249.

Chin-Parker, S., & Bradner, A. (2017). A contrastive account of explanation generation. *Psychonomic Bulletin & Review*, *24*(5), 1387–1397.

Cho, Y. H., & Jonassen, D. H. (2012). Learning by self-explaining causal diagrams in high-school biology. *Asia Pacific Education Review*, *13*(1), 171–184.

Christie, S., & Gentner, D. (2010). Where Hypotheses Come From: Learning New Relations by Structural Alignment. *Journal of Cognition and Development*, *11*(3), 356–373.

Christie, S., & Gentner, D. (2014). Language Helps Children Succeed on a Classic Analogy Task. *Cognitive Science*, *38*(2), 383–397.

Crowley, K., & Siegler, R. S. (1999). Explanation and Generalization in Young Children's Strategy Learning. *Child Development*, *70*(2), 304–316.

Day, S. B., & Goldstone, R. L. (2012). The Import of Knowledge Export: Connecting Findings and Theories of Transfer of Learning. *Educational Psychologist*, *47*(3), 153–176.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Learning by generating vs. receiving instructional explanations: Two approaches to enhance attention cueing in animations. *Computers & Education*, *55*(2), 681–691.

Doumas, L. A. A., & Hummel, J. E. (2013). Comparison and Mapping Facilitate Relation Discovery and Predication. *PLoS ONE*, *8*(6), e63889.

Edwards, B. J., Williams, J., Gentner, D., & Lombrozo, T. (2019). Explanation recruits comparison in a category-learning task. *Cognition*, *185*, 21–38.

Eiriksdottir, E., & Catrambone, R. (2011). Procedural Instructions, Principles, and Examples: How to Structure Instructions for Procedural Tasks to Enhance Performance, Learning, and Transfer. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(6), 749–770.

Fonseca, B. A., & Chi, M. T. H. (2010). Instruction based on self-explanation. In *The handbook of research on learning and instruction*. Routledge.

Forbus, K. D., & Gentner, D. (1986). Learning and physics: Toward a theoretical framework. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2). Kaufmann.

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

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415.

Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. H. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction*, *22*(1), 47–61.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. MIT Press.

Gentner, D. (2010). Bootstrapping the Mind: Analogical Processes and Symbol Systems. *Cognitive Science*, *34*(5), 752–775.

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories: Structure Mapping and Relational Language. *Child Development*, *82*(4), 1173–1188.

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2016). Rapid Learning in a Children's Museum via Analogical Comparison. *Cognitive Science*, *40*(1), 224–240.

Gentner, D., Loewenstein, J., & Thompson, L. (2004). Analogical encoding: Knowledge transfer and integration. In K. D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 452–457). Lawrence Erlbaum Associates, Inc.

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.

Gentner, D., Rattermann, M. J., & Forbus, K., D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, *25*, 524–575.

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, *16*(2), 104–121.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, *137*, 137–153.

Goldwater, M. B., Gentner, D., LaDue, N. D., & Libarkin, J. C. (2021). Analogy Generation in Science Experts and Novices. *Cognitive Science*, *45*(9).

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, *142 7*, 729–757.

Hummel, J. E., Licato, J., & Bringsjord, S. (2014). Analogy, explanation, and proof. *Frontiers in Human Neuroscience*, *8*.

Jamrozik, A., & Gentner, D. (2020). Relational labeling unlocks inert knowledge. *Cognition*, *196*, 104–146.

Kotovsky, L., & Gentner, D. (1996). Comparison and Categorization in the Development of Relational Similarity. *Child Development*, *67*, 2797–2822.

Kubricht, J. R., Lu, H., & Holyoak, K. J. (2017). Individual differences in spontaneous analogical transfer. *Memory & Cognition*, *45*(4), 576–588.

Kurtz, K. J., Miao, C.-H., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, *10*(4), 417–446.

Legare, C. H., & Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *Journal of Experimental Child Psychology*, *126*, 198–212.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315–353.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical Learning in Negotiation Teams: Comparing Cases Promotes Learning and Transfer. *Academy of Management Learning and Education*, *2*(2), 119–127.

Lombrozo, T. (2012). Explanation and Abductive Inference. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*(2), 167–204.

Margulieux, L. E., & Catrambone, R. (2019). Finding the Best Types of Guidance for Constructing Self-Explanations of Subgoals in Programming. *Journal of the Learning Sciences*, *28*(1), 108–151.

Margulieux, L. E., & Catrambone, R. (2021). Scaffolding problem solving with learners' own self explanations of subgoals. *Journal of Computing in Higher Education*, *33*(2), 499–523.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Science*, *25*, 431–467.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory & Cognition*, *19*(6), 543–557.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, *28*(4), 1237–1263.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 510–520.

Rattermann, M. J., & Gentner, D. (1998). More Evidence for a Relational Shift in the Develoment of Analogy: Children's Performance on a Causal-mapping Task. *Child Development*, *13*, 453–478.

Richey, J. E., Zepeda, C., D., & Nokes-Malach, T. J. (2015). Transfer effects of prompted and self-reported analogical comparison and self-explanation. In D. Noelle C., R. Dale, A. Warlaumont S., J. Yoshimi, T. Matlock, C. Jennings D., & P. Maglio P. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1985–1990). Cognitive Science Society.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). MATHEMATICS: Cognitive Supports for Analogies in the Mathematics Classroom. *Science*, *316*(5828), 1128–1129.

Rittle-Johnson, B., Fyfe, E. R., Loehr, A. M., & Miller, M. R. (2015). Beyond numeracy in preschool: Adding patterns to the equation. *Early Childhood Research Quarterly*, *31*, 101–112.

Rittle-Johnson, B., & Loehr, A. M. (2017). Eliciting explanations: Constraints on when self-explanation aids learning. *Psychonomic Bulletin & Review*, *24*(5), 1501–1510.

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*(3), 529–544.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 456–468.

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.

Roy, M., & Chi, M. T. H. (2005). The Self-Explanation Principle in Multimedia Learning. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (1st ed.). Cambridge University Press.

Sidney, P. G., Hattikudur, S., & Alibali, M. W. (2015). How do contrasting cases and self-explanation promote learning? Evidence from fraction division. *Learning and Instruction*, *40*, 29–38.

Snoddy, S., & Kurtz, K. J. (2017). Promoting Spontaneous Analogical Transfer: The Role of Category Status. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar J. (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 3215–3220). Cognitive Science Society.

Son, J. Y., Doumas, L. A. A., & Goldstone, R. L. (2010). When do words promote analogical transfer? *The Journal of Problem Solving*, *3*(1), 52–92.

Son, J. Y., & Goldstone, R. L. (2009). Contextualization in Perspective. *Cognition and Instruction*, *27*(1), 51–89.

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483.

Thompson, C., & Opfer, J. (2010). How 15 Hundred Is Like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition. *Child Development*, *81*(6), 1768–1786.

Walker, C. M., Lombrozo, T., Legare, C. H., & Gopnik, A. (2014). Explaining prompts children to privilege inductively rich properties. *Cognition*, *133*(2), 343–357.

Williams, J. J., & Lombrozo, T. (2010). The Role of Explanation in Discovery and Generalization: Evidence From Category Learning. *Cognitive Science*, *34*(5), 776–806.