Scaffolding to Support Analogical Comparisons with Science Images

Benjamin D. Jee[1], Florencia K. Anggoro[2], Andja Kola[1], Bryan J. Matlen[3], and Dedre Gentner[4]

[1]Department of Psychology, Worcester State University
[2]Department of Psychology, College of the Holy Cross
[3]Learning and Technology Program, WestEd
[4]Department of Psychology, Northwestern University

Abstract

This study examines scaffolding techniques to support students' analogical comparisons with science images. Analogical comparison involves identifying deep relational structure over superficial similarities, which can be challenging without guidance. Across experiments, participants compared analogous evolutionary science images with or without various scaffolds: describing the relation, completing a mapping table, and spatial support for alignment. Results demonstrated that describing the relation and completing a mapping table, especially in combination, significantly enhanced scientific interpretations of the images. However, spatial support was ineffective. The findings highlight practical strategies for improving conceptual learning from science visuals.

Keywords: Analogical comparison; scaffolding; science comprehension; visual representations

Introduction

When Charles Darwin encountered the Galapagos Islands' finches, he noted the remarkable diversity of their beaks and how their dimensions fit the foods available in different habitats, like seeds or fruits. The origin of such patterns was a matter of fierce debate, and Darwin drew further comparisons to lend support to his evolutionary explanation. He recognized that Galapagos finches were more like South American species than to birds found on islands with a similar climate, vegetation, and terrain as the Galapagos. Darwin took the relation between continental and island birds as evidence of a shared ancestry and proposed that Galapagos finches originated from America before adapting to different island ecologies through *natural selection*: individuals with beaks suited to an available food source were more likely to survive and pass on their traits to future generations.

The relational reasoning that Darwin exemplified—aligning observable examples to uncover deeper structures—remains central to science learning today. Indeed, many scientific concepts are defined not by overt features but by relational structure (Goldwater & Schalk, 2016). Take the concept of *fitness* that is central to evolutionary theory. Fitness is not an inherent quality of an organism; it depends on whether its traits confer survival and reproductive advantages over other individuals or groups. Moreover, traits that increase fitness in one setting, like the Arctic, can be maladaptive in another, like a tropical rainforest. Understanding fitness requires thinking beyond an organism's visible appearance, focusing instead on how its traits function in a broader ecological context.

Visual representations play a vital role in conveying relational concepts in science (Jee et al., 2022; Mayer, 2020). Consider images A and B in Figure 1, which illustrate the concept of evolutionary fitness with two species. Image A shows a Galapagos tortoise with a saddle-shaped shell and long neck (bottom right) able to feed on a tall cactus plant whereas tortoises with domed shells and shorter necks are unable to do so. In this hot, dry environment, long-necked tortoises have a survival advantage because they can access a scarce food resource. Image B shows peppered moths clinging to a tree darkened by factory pollution. A moth with dark coloring (bottom left) blends into the bark and goes unseen by predators, such as birds. Light-colored moths stand out and are more vulnerable. In this setting, dark moths have a survival advantage. The scenarios in Figure 1 thus share a deep relational structure, showing how a trait (long neck or dark color) improves the fitness of some individuals, while those without the trait are less likely to survive.

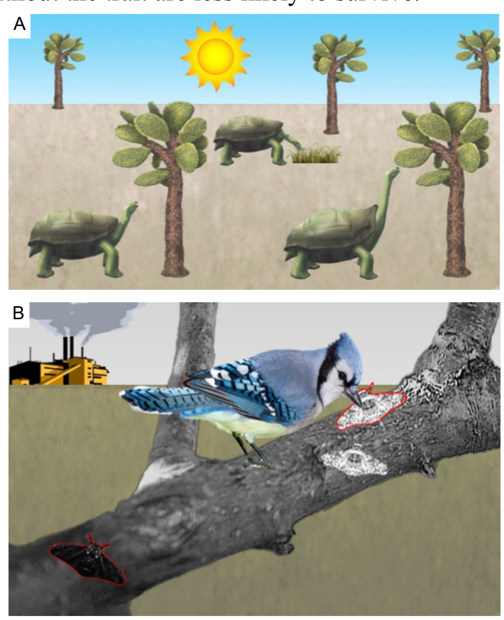

Figure 1: Scenes illustrating Galapagos tortoise (A) and peppered moth (B) evolution.

To grasp the concept of fitness expressed in Figure 1, students must represent the organisms in terms of their role in a system of relations, and resist matching them using more superficial similarities. Students with little prior knowledge can have difficulty identifying these deeper connections (e.g., Chi et al., 1981). However, relational structure can become salient when students engage, as Darwin himself did, in an explicit comparison of the examples (Goldwater & Gentner, 2015; Jee et al., 2013). By comparing the two analogous scenes in Figure 1, a student may gain a more robust

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

understanding of fitness in general. The purpose of the present research was **to test scaffolding methods to support analogical comparisons with science images**.

Analogical comparison involves a process of structural alignment in which the elements of two examples are matched on the basis of their role in a common system of relations (Gentner, 2010). For Figure 1, the long-necked tortoise in Image A should be matched with the dark moth in Image B because both animals possess traits that increase their evolutionary fitness. Yet, students may not apprehend this deeper relation without some external support, instead focusing on shallow similarities and differences (Markman & Gentner, 1993; Kurtz et al., 2001). With the scenes in Figure 1, a shallow comparison could lead a student to match the long-necked tortoise with the bird in Image B, because both are eating. While not insensible, this mapping is ultimately inaccurate and incomplete, as it does not produce a coherent match for the other elements in the scene, such as the short-necked tortoise in Image A. More broadly, the superficial matching fails to reflect the evolutionary concepts that link the scenes together. If students merely skim over the images during comparison, they could fail to grasp the scientific content and come away with a superficial interpretation.

Several forms of scaffolding have been used in prior studies of analogical learning and instruction. We focus on three that seem especially promising for comparisons involving the kinds of visual representations that appear in science textbooks and other educational media. The first is to prompt the student to explicitly describe and/or explain the common relation between the images. This task orients the student toward relational commonalities as opposed to surface similarities (Gick & Holyoak, 1983; Kurtz et al., 2001). For example, in a study of space science learning, elementary school students were asked to explain the link between events, like sunrise, as seen from the familiar Earth-based perspective and from a space-based perspective, with the full planet in view (Jee & Anggoro, 2019). Students who received this scaffolding learned more about the causal link between Earth's rotation and the Sun's apparent motion than those who did not. Prompting students to explain the relation between the scenes in Figure 1 could have a similar benefit, drawing attention to the causal role of each object in an evolutionary system.

The second scaffolding technique involves completing a mapping table to specify the corresponding elements in two images. Filling out a partially-completed mapping table—where the elements of only one example are provided—can help students identify common relational structure and reason about the cases (Kurtz et al., 2001; Mason, 2004). This task could promote systematic comparison by discouraging incoherent sets of correspondences. For example, if students complete a mapping table for the images in Figure 1, they could realize that matching the long-necked tortoise with the bird leads to a dead end, as no coherent match would remain for the short-necked tortoise.

The third scaffolding technique involves modifying the spatial layouts of images to facilitate alignment and reduce interference from irrelevant or competing matches. This is achieved by standardizing the internal layouts of two images so that key elements share the same relative positions and then placing the images above or beside each other to ensure the most direct, unimpeded pathways between corresponding parts. Spatially aligning simple images improves children's and adults' speed and accuracy in detecting similarities and differences, including abstract relations (Matlen et al., 2020; Zheng et al., 2022). Spatial alignment can also facilitate comparison of visually-complex representations, like skeletal structures, helping students detect mismatched elements (Simms et al., 2023). In a spatially-aligned version of Figure 1, for example, the long-necked tortoise would occupy the bottom left, matching the dark moth's position (see Fig. 3 below). This spatial alignment could help students identify the overlapping conceptual connections between the images.

Our aim was to evaluate the effects of these scaffolding techniques without instructor guidance or feedback. This mirrors the experience of students learning from textbooks and other media, where they must independently interpret complex visual representations. Participants in our studies compared two analogous science images and were asked to match a key element between them. Our experimental conditions involved one or more forms of scaffolding during comparison, while our control condition involved visual comparison without scaffolding.

Experiment 1

Experiment 1 tested whether scaffolded comparison helps students detect scientifically-relevant connections between the animals in Figure 1 (Images A and B) more effectively than unscaffolded comparison. The Scaffolded condition included two tasks: describing/explaining the relation between the scenes and completing a mapping table to identify a match in Image B for the long-necked tortoise, short-necked tortoise, and tall cactus plant. In the Unscaffolded condition, participants identified a match in Image B for only the long-necked tortoise.

Our primary measure was whether participants matched the long-necked tortoise to its scientifically-relevant counterpart: the dark moth. All participants also completed two interpretation items about the images: listing two differences between the scenes and rating the similarity of the scenes. They were also given two reasoning questions: one about the peppered moths (inferring what would happen if the events in Image B repeated over a long period) and one about a new species (explaining how pelicans evolved from short-beaked to long-beaked forms). If support for comparison promotes scientifically-accurate interpretations, then participants in the Scaffolded condition should match the long-necked tortoise with the dark moth more often, focus on relational differences between the scenes, and judge the scenarios as more similar. Furthermore, if scaffolding supports evolutionary reasoning, these participants should perform better on the reasoning questions as well.

Method

Participants. One hundred five undergraduates (82% female) from two liberal arts colleges in the Northeastern U.S. (one private, one public) participated in person for course credit. Their mean age was 19.5 years (*SD* = 2.6). Participants were randomly assigned to either the Scaffolded or Unscaffolded comparison condition.

Materials. The primary materials consisted of the two educational images on evolutionary fitness shown in Figure 1. The research team created the images based on diagrams in middle school science textbooks.

Procedure. All instructions and tasks were administered via Qualtrics on a computer in a research lab. Participants first provided consent and were informed that they would see two images from a lesson on evolution: one showing a species of tortoise from the Galapagos Islands and the other a species of moth from Northern England. They were then shown images A and B as in Figure 1. Participants in the Scaffolded condition were asked to compare the two scenarios and type a description of what is happening in the images and why. Next, they completed a mapping table, identifying objects in the moth image corresponding to the long-necked tortoise, short-necked tortoise, and tall cactus plant. Participants in the Unscaffolded condition viewed the same images but were asked to identify the match for only the long-necked tortoise. After the initial comparison, all participants completed the following tasks with the images visibale on screen: 1) list two differences between the scenes, 2) rate the similarity of the scenarios from 0 (very low) to 100 (very high), 3) answer a multiple-choice question about how the moth species would evolve if the events in the image repeated over a long period (correct answer: the species would be mostly dark-colored), and 4) answer an open-ended question about pelican evolution: "Pelicans have large beaks that enable them to scoop fish out of the water and gulp them down. The ancestors of pelicans—who lived long, long ago—did not have large beaks. How do you think pelicans came to have large beaks?" (based on Shtulman et al., 2016). Finally, participants indicated the number of high school and college biology courses they had taken.

Results

Data coding procedures for this and the other experiments involved multiple coders who achieved a criterion reliability level of $K \geq .80$ before coding independently. Our primary measure of whether scaffolding affected participants' interpretation of the images' deeper scientific meaning was accuracy on the mapping task—specifically, whether participants matched the long-necked tortoise to the dark moth, as both possess traits that enhance fitness. As shown in Figure 2, participants in the Scaffolded condition made the scientifically-relevant match significantly more often (*M* = .57, *SD* = .50) than those in the Unscaffolded condition (*M* = .27, *SD* = .45), $t(103) = 3.23$, $p < .001$, $d = .63$. This effect persisted when controlling for participants' high school and college biology experience, which had only a weak association with matching accuracy, $r(105) = .11$, $p = .27$.

To further probe these findings, we coded whether participants in the Scaffolded condition mentioned key concepts, such as fitness, adaptation, and survival, in their written descriptions/explanations during the comparison phase. Overall, 54% of the participants in the Scaffolded condition mentioned one or more of these concepts in their responses. Those who did were significantly more likely to respond with the scientifically-relevant match (*M* = .76, *SD* = .44) than those who did not (*M* = .36, *SD* = .46), $t(52) = 3.14$, $p < .01$, $d = .86$. Thus, scaffolded comparison, especially when participants activated relevant scientific concepts, was strongly associated with scientifically-relevant interpretations of the images

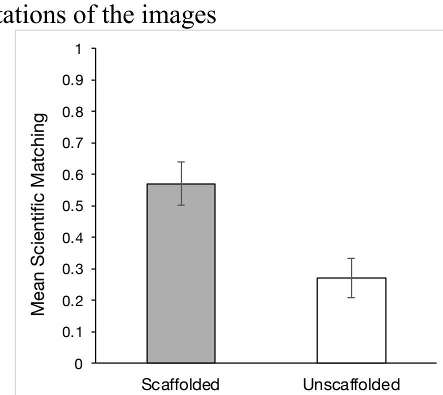

Figure 2: Mean scientific matches (+/- SEM) by Exp. 1 condition

To assess whether participants focused on relational structure when comparing the images, we coded whether they listed differences involving relations (e.g., "Image A's evolutionary trait focuses on getting a resource and image B's focuses on staying hidden from a predator") or not (e.g., "There is a factory in image B"). Participants in the Scaffolded condition listed more relations (*M* = 0.69, *SD* = 0.67) than those in the Unscaffolded condition (*M* = 0.55, *SD* = 0.73), though the difference was not significant, $t(103) = 1.00$, $p = .16$, $d = .20$. However, participants in the Scaffolded condition rated the scenes as significantly more similar (*M* = 54.3, *SD* = 24) than those in the Unscaffolded condition (*M* = 46.1, *SD* = 21.9), $t(101) = 1.80$, $p = .04$, $d = .35$.

The scaffolding manipulation had no discernable effect on participants' scientific reasoning. For the inference question about the moth species, accuracy was comparably high for the Scaffolded (*M* = .76, *SD* = .43) and Unscaffolded condition (*M* = .80, *SD* = .40), $t(103) = 0.55$, $p = .29$, $d = -.11$. For the question about pelican evolution, which we assessed for the use of five key scientific concepts (variation, inheritance, survival, reproduction, and population change; based on Shtulman et al., 2016), participants in the Scaffolded mentioned about the same number of concepts (*M* = 1.1, *SD* = 1.6) as those in the Unscaffolded condition (*M* = 1.0, *SD* = 1.6), $t(103) = 0.05$, $p = .48$, $d = .01$.

Discussion

The results demonstrate that scaffolded comparison helped participants identify a scientifically-relevant match between the images. Scaffolded participants who expressed key scientific concepts (e.g., fitness, adaptation, survival) during comparison were particularly likely to identify the correct match. The Scaffolded condition's trend toward listing more relational differences and their higher similarity ratings for the two scenes provides further evidence that scaffolding drew attention to the deeper, evolutionary relations that unite the two scenes. However, we did not find scaffolding effects on the scientific reasoning tasks, suggesting that the benefits of scaffolding may not extend beyond interpreting the visually available information.

Experiment 2

Experiment 2 builds on Experiment 1 by investigating the effects of the Scaffolded comparison condition for spatially aligned images, in which corresponding elements are directly adjacent in space (Matlen et al., 2020). Experiment 2 involved the same Scaffolded and Unscaffolded conditions as Experiment 1, but we redesigned the tortoise evolution image to place the long-necked tortoise in the bottom left (mirroring the location as the dark moth) and the short-necked tortoise on the right (like the light moth). We also moved the Sun to match the factory's relative location, as both play a role in driving environmental change. The resulting spatially-aligned pair is shown in Figure 3.

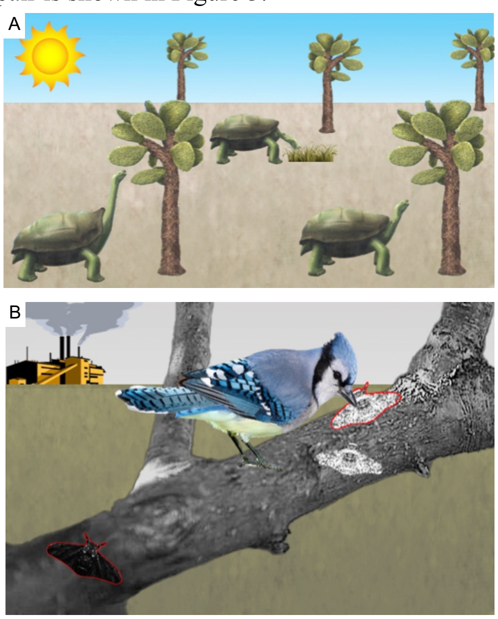

Figure 3: A spatially-aligned version of the images in Figure 1.

If spatial alignment facilitates scientific interpretation of the images, additional scaffolding may be unnecessary. If so, participants in the Unscaffolded condition should perform comparably to those in the Scaffolded condition. This would fit with prior findings that spatial supports improve alignment (Matlen et al., 2020; Zheng et al., 2022). Alternatively, spatial alignment may enhance performance across all participants, raising baseline performance without affecting the relative advantage for the Scaffolded group. It is also possible that the spatial adjustments will have little impact, resulting in a similar scaffolding effect to that seen in Experiment 1.

Method

Participants. Eighty-nine undergraduates (81% female) from the same two liberal arts colleges in the Northeastern U.S. participated in person for course credit. Their mean age was 20.0 years (*SD* = 3.0). Participants were randomly assigned to either the Scaffolded Comparison condition or the Unscaffolded Comparison condition.

Materials and Procedure. The Scaffolded and Unscaffolded Comparison conditions were the same as in Experiment 1, as were the post-comparison measures: listing two differences between the images, rating similarity, inferring how the moth species would evolve, and explaining pelican evolution. The only difference was that participants were presented with the spatially-aligned images shown in Figure 3.

Results

Our main question concerns how Scaffolded vs. Unscaffolded comparison influences scientific interpretation for images that include spatial support for alignment. As shown in Figure 4, participants in the Scaffolded condition were significantly more likely to match the long-necked tortoise with the dark moth (M = .58, SD = .50) than those in the Unscaffolded condition (M = .16, SD = .37), $t(87)$ = 4.50, p < .001, d = .95. As in Experiment 1, this effect held when controlling for participants' biology course experience, which was not significantly related to mapping accuracy, $r(88)$ = -.08, p = .45. We also found that 53% of participants in the Scaffolded condition mentioned relevant evolutionary concepts in their written descriptions/explanations of the scenes, and those who did were more likely to make the relevant match (M = .79, SD = .42) than those who did not (M = .33, SD = .48), $t(43)$ = 3.43, p < .001, d = 1.02. Overall, the scaffolding effects with spatially-aligned images were similar to the effects with non-aligned images in Experiment 1.

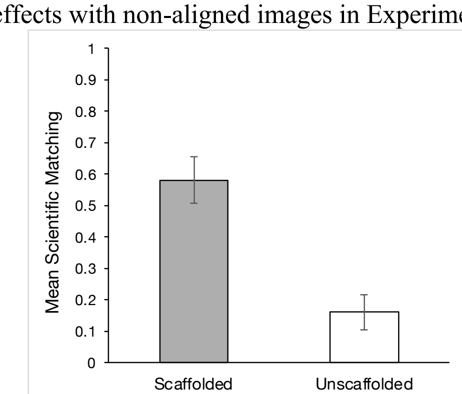

Figure 4: Mean scientific matches (+/- SEM) by Exp. 2 condition

In the difference-listing task, there was a nonsignificant trend such that participants in the Scaffolded condition responded with slightly more relations (M = 0.67, SD = 0.83) than those in the Unscaffolded condition (M = 0.52, SD = 0.73), , $t(87)$ = .87, p =.19, d = .18. Participants in the Scaffolded condition rated the scenes as more similar (M = 48.5, SD = 22.0) than those in the Unscaffolded condition (M = 42.3, SD = 19.1), but this trend was also nonsignificant, $t(85)$ = 1.41, p =.08, d = .30.

Unlike Experiment 1, participants in the Scaffolded condition were significantly more accurate at inferring how the moth species would evolve (M = .82, SD = .39) than those in the Unscaffolded condition (M = .64, SD = .49), $t(87)$ = 2.00, p =.02, d = .42. However, explanations of pelican evolution continued to lack the key concepts we coded for, with participants in the Scaffolded condition mentioning an average of 1.1 concepts (SD = 1.6) and those in the Unscaffolded condition mentioning 1.0 (SD = 1.3), $t(87)$ = 0.36, p =.36 d = .08.

Discussion

The results of Experiment 2 were largely consistent with Experiment 1, underscoring the benefits of scaffolded comparison for students' scientific interpretation of visual representations. The spatial alignment of the images had no discernable effect on participants' performance. Thus, spatial support alone was not sufficient for helping participants in the Unscaffolded condition identify the scientifically-relevant correspondence at the level of participants in the Scaffolded condition, nor did it facilitate performance on the other measures of structural alignment and scientific reasoning. One notable difference in the results was that participants in the Scaffolded condition performed more accurately on the inference question about moth evolution, suggesting that scaffolded comparison may support not only interpretation of but also reasoning about the depicted scenes.

Experiment 3

The previous studies indicate that scaffolded comparison enhances students' comprehension of analogous visual representations in science by helping them identify deeper structural connections. Students who engaged in unscaffolded comparison were less likely to form a scientific interpretation and identify a relevant conceptual link, and this pattern held even when the images were spatially arranged to promote structural alignment. We designed Experiment 3 to explore the source of the scaffolding effect in more depth. In our Scaffolded comparison condition, participants completed two tasks: 1) writing a description of what is happening and why, and 2) filling out a three-part mapping table to match the long-necked tortoise, short-necked tortoise, and tall cactus plant. While both tasks are thought to promote structural alignment, it is possible that only one—or one in particular—is essential for identifying the relevant connection between the images. To test this, we manipulated the scaffolding tasks between groups in Experiment 3, with one group completing only a written description, and another completing only a three-part mapping table.

Method

Participants. The study was conducted online and remotely using the same undergraduate population as the previous studies. Anticipating greater within-group variability due to the uncontrolled online setting and a smaller between-groups effect (as both conditions involved scaffolding), we increased the sample size to approximately double that of the previous experiments. A total of 186 undergraduates (85% female) from a public liberal arts college in the Northeastern U.S. participated for course credit. We eliminated 12 participants who took more than 67 minutes to complete the study, which exceeds the maximum time taken by any participant from Experiment 2. This left a final sample of 174 participants (mean age = 21.4 years, SD = 6.4). Participants were randomly assigned to either the Describe Relations condition or the Mapping Table condition.

Materials and Procedure. All participants were shown the spatially-aligned images from Experiment 2 (Figure 3). Participants in the Describe Relations condition completed the written description task from prior experiments, describing what was happening and why. They were then asked to identify the match for the long-necked tortoise. In the Mapping Table condition, participants completed only the 3-part mapping task, identifying matches for the long-necked tortoise, short-necked tortoise, and tall cactus plant. The measures and procedure were otherwise the same as those used in the prior experiments.

Results

As in the prior studies, our main interest was whether participants matched the long-necked tortoise with the dark-colored moth, the scientifically-relevant connection. As shown in Figure 5, participants in the Mapping Table condition made this match more often (M = .47, SD = .50) than those in the Describe Relations condition (M = .36, SD = .48), though the difference was not significant, $t(172)$ = 1.44, p = .15, d = .22. This trend remained when we accounted for biology course experience, which was not significantly related to mapping accuracy, $r(174)$ = .08, p = .30. In the Describe Relations condition, 46% of participants mentioned relevant evolution concepts. Those who did were significantly more likely to make the scientific match (M = .52, SD = .51) than those who did not (M = .22, SD = .42), $t(89)$ = 3.08, p < .01, d = .65. So, the benefits of describing relations depended on whether the participant applied relevant scientific concepts during the comparison.

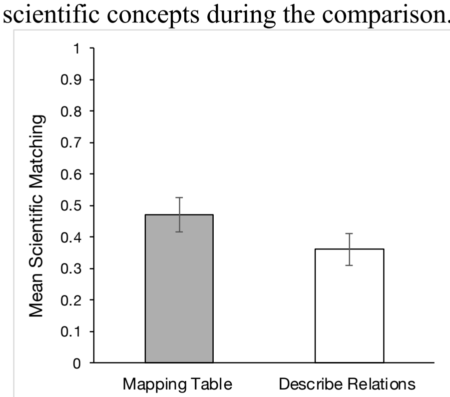

Figure 5: Mean scientific matches (+/- SEM) by Exp. 3 condition

On the post-comparison measures, it was participants in the Describe Relations condition who favored structural relations, though the between-groups differences were not statistically significant. When listing differences, they mentioned relations slightly more often (M = 0.45, SD = 0.67) than those in the Mapping Table condition (*M* = 0.41, *SD* = 0.72), *t*(172) = 0.40, *p* =.70, *d* = .06. They also rated the similarity of the images as slightly higher, *M* = 45.8 (*SD* = 21.0) vs. *M* = 42.1 (*SD* = 22.2), *t*(172) = 1.11, *p* =.27, *d* = .17. Participants in the Describe Relations condition also responded somewhat more accurately when inferring how the moth species would evolve (*M* = .76, *SD* = .43) than those in the Mapping Table condition (*M* = .67, *SD* = .47), *t*(172) = 1.22, *p* =.22, *d* = .18. However, those in Mapping Table condition mentioned slightly more key concepts in their explanations of pelican evolution (*M* = 1.33, *SD* = 1.29) than participants in the Describe Relations condition (*M* = 1.15, *SD* = 1.08), *t*(172) = 0.95, *p* =.34, *d* = .15.

We conducted an additional study-wide analysis to determine whether the benefits of the two effective scaffolding methods—describing relations and completing a mapping task—are additive or redundant. We pooled data across the studies (*N* = 369) and sorted participants into three groups based on whether they completed 0, 1, or both scaffolding tasks, collapsing across spatial alignment. The means for scientific matching are shown in Figure 6. The level of scaffolding had a significant effect, $F(2, 366) = 13.77$, $< .001$, $\eta_p^2 = .07$. Post hoc tests with Bonferroni correction found that participants who performed both scaffolding tasks had higher scientific matching than those who completed only one ($p = .02$), who in turn had higher matching than those who performed neither task ($p < .01$).

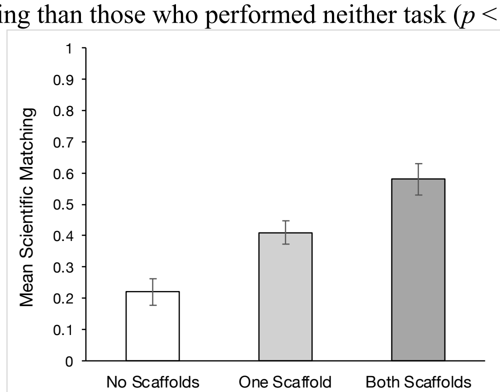

Figure 6: Mean scientific matches (+/- SEM) by level of scaffolding

Discussion

Experiment 3 revealed a nonsignificant trend such that participants who completed a full mapping task exhibited more scientific matching than those who completed the description/explanation task. While this could signify deeper relational thinking, it is worth noting that the dependent measure overlapped with the scaffolding task in the Mapping table condition, possibly giving that group an advantage. Looking at the post-comparison measures, participants in each condition performed similarly, with slight differences favoring the Describe Relations condition. Importantly, our study-wide analysis indicated that participants completing both tasks were most likely to make scientifically relevant matches, suggesting complementary rather than redundant benefits, and highlighting the advantage of combining tasks to enhance understanding of structural relations.

General Discussion

As Darwin demonstrated long ago, comparison is a powerful tool for fostering scientific insights. In the present experiments, we explored how scaffolding comparison—via prompts to describe relations and complete a mapping table—supports students in interpreting analogous science images. Across experiments, scaffolding promoted structural-scientific interpretation of the images, with participants who engaged in both scaffolding tasks showing the strongest inclination toward relational responses. Importantly, these effects were achieved without instructor involvement or feedback, suggesting that these scaffolding tasks are effective even in unsupervised learning contexts.

Interestingly, the spatial alignment of images, which has been shown to optimize comparison in prior research (e.g., Matlen et al., 2020), did not significantly alter performance in our study. This suggests limits to spatial alignment effects. From prior findings, we suspect that spatial alignment may be most effective when the relationships being evaluated are inherently spatial; for example, spatial alignment facilitates comparisons of skeletal structures, where the main considerations were the size, shape, and location of individual bones (Simms et al., 2023). In contrast, a complex science concept like evolutionary fitness, even when represented diagrammatically, is not inherently spatial, so may be less compatible with spatial supports. Nevertheless, spatial alignment could still support reasoning in contexts where spatial and other conceptual relationships must be integrated, such as learning how observable phenomena, like thermal expansion, are related to underlying mechanisms, like molecular activity (Snir et al., 2003). Further research is needed to clarify when spatial alignment is most useful.

The present study contributes to a growing body of evidence that scaffolding techniques can optimize learning from comparison across different contexts. Teachers can support students' scientific thinking by incorporating scaffolding tasks like describing relations and completing mapping tables, alongside other cognitive supports like verbal prompts. relational language, and linking gestures (Richland & McDonough, 2009; Yuan et al., 2017). These methods help students focus on structural relationships and activate relevant concepts, potentially amplifying the effects of other scaffolding while promoting deeper understanding.

To further support science learning and instruction, future research should explore how different scaffolding techniques apply—both individually and in concert—across diverse contexts and learners. For example, less-experienced students or those encountering unfamiliar concepts students may require more support as they compare novel information, while experienced learners may benefit from more open-ended scaffolds. Additionally, studies should investigate how scaffolding influences various stages of analogical comparison, including identifying correspondences, reasoning through relationships, and generalizing knowledge. By addressing these questions, we can better understand how to tailor scaffolding strategies to meet the needs of learners across a wide range of educational settings.

Acknowledgments

We thank Hiba Farhan and Louisa Wetzel for assistance with data collection and coding. The research reported here was supported by NSF #2201894, #2201895, #2201896, and #2201897, and by IES #R305A170165. The opinions expressed are those of the authors and do not necessarily represent views of the funding agencies.

References

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121–152.

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

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

Jee, B. D., & Anggoro, F. K. (2019). Relational scaffolding enhances children's understanding of scientific models. *Psychological Science, 30*(9), 1287-1302.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C. A., Shipley, T. F., & Sageman, B. (2013). Finding faults: Analogical comparison supports spatial concept learning in geoscience. *Cognitive Processing, 14*(2), 175–187. https://doi.org/10.1007/s10339-013-0551-7

Jee, B. D., Matlen, B. J., Greenlaw, M., Simms, N., & Gentner, D. (2022). Spatial supports for comparison in educational science images. *Instructional Science*, *50*(6), 807-827.

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

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*(4), 431–467.

Mason, L. (2004). Fostering understanding by structural alignment as a route to analogical reasoning. *Instructional Science, 32*(3), 293-318.

Matlen, B. J., Gentner, D., & Franconeri, S. L. (2020). Spatial alignment facilitates visual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 46*(5), 443–457.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology, 35*(1), 28–43.

Shtulman, A., Neal, C., & Lindquist, G. (2016). Children's ability to learn evolutionary explanations for biological adaptation. *Early Education and Development, 27*(8), 1222–1236. https://doi.org/10.1080/10409289.2016.1154418

Simms, N. K., Matlen, B. J., Jee, B. D., & Gentner, D. (2023). Spatial alignment supports comparison of life science images. *Journal of Experimental Psychology: Applied*, *29*(4), 747–760. https://doi.org/10.1037/xap0000471

Snir, J., Smith, C. L., & Raz, G. (2003). Linking phenomena with competing underlying models: A software tool for introducing students to the particulate model of matter. *Science Education, 87*(6), 794-830.

Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. *Developmental Psychology, 53*(6), 1098–1114. https://doi.org/10.1037/dev0000302

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