Do Analogies in Geoscience Textbooks Inflate Judgments of Understanding?

Thomas D. Griffin (tgriffin@uic.edu)

Lena Hildenbrand (lhilde3@uic.edu)

Jennifer Wiley (jwiley@uic.edu)

University of Illinois at Chicago, Department of Psychology
1007 West Harrison Street, Chicago IL 60607 USA

Abstract

Analogies are widely used in science textbooks to connect unfamiliar phenomena to more familiar everyday objects and processes in the hope they will help students better understand the scientific explanations of complex processes. However, the emerging literature on the *illusion of explanatory depth* suggests that people are highly overconfident in their understanding of the causal processes underlying how everyday objects and devices work. Thus, analogies involving familiar everyday objects may inflate judgments of understanding (JOUs) for the to-be-learned scientific explanations. Two experiments tested whether common analogies used in geology textbooks increase students' JOUs for geology phenomena. A first experiment showed that analogies notably increase JOUs when presented and judged as individual sentences. A second experiment showed that when these same sentences containing analogies were embedded within longer explanatory texts, they did not help (and sometimes harmed) actual understanding. Readers showed general overconfidence. However, unlike the first experiment, when readers made global JOUs about their overall understanding of entire texts, those JOUs did not differ depending on whether there were analogies embedded within the texts. Results extend prior work on the effect of analogies on metacomprehension, and have potential implications for the interpretation of findings related to the illusion of explanatory depth.

Keywords: Metacomprehension; Analogies; Illusion of Explanatory Depth; Geoscience Instruction; Judgments of Understanding

Introduction

Science textbooks about natural phenomena often convey complex processes about phenomena, including detailed causal explanations containing specific jargon terms, such as "When a glacier melts, lateral moraines lie stranded along the side of the glacially carved valley." These very specific aspects of geological processes described with esoteric jargon are unlikely to seem particularly familiar to students. However, science textbooks often employ the use of analogies where the complex phenomena are said to be like some more familiar everyday object or event (Curtis & Reigeluth, 1984; Thiele & Treagust, 1994).

Analogies are employed in textbooks because of their theorized and empirically demonstrated benefit for learning due to allowing a student to apply what they know about the source object to the less familiar target object (Donnelly & McDaniel, 1993; Gentner, 1983). While well-constructed analogies could help students to understand a geological explanation more deeply, the introduction of a familiar everyday object also introduces an opportunity for greater metacognitive overconfidence. The student may perceive a boost of confidence in understanding from the mention of the source object, but this is likely to reflect an illusion unless the analogy actually improves understanding. If not properly processed in a way that fosters transfer of knowledge about the source object to the science concept, analogies might only prompt the transfer of the sense of familiarity and thus an increase in metacognitive JOUs about the science concepts, resulting in greater overconfidence. Metacomprehension research shows that readers are consistently overconfident in their understanding of new scientific explanations they have just read (e.g., Griffin et al., 2025; Hildenbrand et al., 2024), and analogies could cause further inflation.

This potential for analogies to lead to inflated JOUs is bolstered by the findings about *illusions of explanatory depth* (IOED). The IOED is when people are highly overconfident in the understanding of how everyday objects and devices work (e.g., toilets, cell phones, sewing machines) as first shown by Rozenblit and Keil (2002). IOED research differs from metacomprehension in that it focuses on overconfidence in understanding the causal processes underlying common devices that people are already familiar with. In a set of 12 studies, Rozenblit and Keil presented participants with names of various devices and asked them to rate "how well you feel you understand how each device works". Then, participants were asked to write an explanation of how that device works and re-rate their understanding. The sizable decline in JOUs from prior to after writing an explanation was taken as evidence that initial JOUs are inflated above actual understanding. Those JOUs decline further after answering a question requiring causal understanding of the mechanisms.

These seminal studies on IOED showed that, besides familiarity with the devices, another strong bivariate predictor of JOUs and overconfidence was how visible that parts of the object were. This suggested that people's ability to visualize and mentally animate the phenomena increases their JOUs. This is important because many analogies used in geology textbooks employ source objects that are both more familiar to learners and more directly visible to learners than the geology phenomena they are being associated with. For

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

example, *a boat rising and sinking in the water as cargo is added or removed from it* is more visible than *the lowering or raising of large areas of lithospheric crust in the asthenosphere over thousands of years as mass of rock is added or subtracted from the lithosphere*. The enormous physical and time scales of these geological events make them not directly observable. Thus, multiple aspects of source objects/phenomena used in analogies for geology are likely to evoke a feeling of understanding that could wind up being transferred onto the geology explanations themselves.

Griffin et al. (2024) examined 28 analogies from a widely used college-level geoscience textbook. They separated the parts of the sentences that referenced the source objects from the parts that referenced the target geology concepts. Descriptions of the source objects had higher concreteness and imageability metrics than descriptions of the geology concepts, according to the MRC Psycholinguistic Database (Coltheart, 1981). Also, students rated the source objects as much more familiar than the geology concepts (Cohen's d = 1.66). Thus, the source objects in commonly used geoscience analogies introduce multiple properties that are also common to everyday devices assumed to be responsible for illusions of understanding about those devices.

In addition, although the IOED refers to JOUs for already known devices, theoretical accounts of the IOED judgment process share a key property with the situation-model perspective on metacomprehension (for a review, see Griffin, Mielicki, & Wiley, 2019). Both accounts assume that people confuse different levels of understanding, mistaking a more superficial description for a deeper understanding of actual causal mechanisms. There is a greater overconfidence in JOUs for explanations about how everyday devices work compared to JOUs for retrieving facts or describing sequence of procedural steps (Rozenblit & Keil, 2002). Likewise, JOUs for scientific texts more accurately predict test performance when the understanding being assessed requires mere memory for explicitly stated facts versus being able to answer inference questions about the complex causal relationships implied by the text (Griffin, Wiley, & Thiede, 2019). In both literatures, people's confusion about different levels of understanding is partially explained by people's lack of experience with causal explanations, whether having to generate explanations about everyday devices or having to answer inference questions that require comprehension of causal mechanisms of scientific phenomena.

The transfer of a feeling of familiarity from a source object to the target concepts in an analogy is likely to reflect problematic overconfidence and illusion of understanding, if useful causal knowledge about the source objects are not also transferred to knowledge of the target geology concepts. Such analogical transfer of knowledge typically requires that the analogy be accompanied by an in-depth explanation in terms of how the elements of the source object map onto the science concept (Jee et al., 2010). A review of how analogies are used in geology textbooks shows these key facilitators of knowledge transfer are rarely present. An analysis of 7 widely used textbooks and online instructional materials for Introductory Geoscience identified 874 analogies, and 81% of them did not include mappings between the source object and geology concept (Poe et al., 2024). Most often the geology phenomena are merely stated to be like some more common object that students are presumed to be more familiar with. A similar problem has been identified with analogy use in textbooks for other sciences (for a review, see Orgill & Bodner, 2006). Superficially presented analogies may act like seductive images often used in textbooks, which create a false sense of familiarity or fluency while not being particularly helpful, thereby inflating JOUs and overconfidence (Serra & Dunlosky, 2010; Wiley, 2019).

Even without analogies, students are generally overconfident in their level of understanding of texts they read about scientific explanations (Dunlosky & Rawson, 2012; Fischchoff, Slovic, & Lichtenstein, 1977; Hacker et al., 2000). The present concern is whether the greater familiarity of the source objects in analogies can boost this confidence even higher. The two prior studies that tested for effects of analogies on JOUs about scientific phenomena failed to observe a general increase in JOUs (Jaeger & Wiley, 2015; Wiley et al., 2018). However, those studies had participants judge their understanding of full texts that explained the phenomena and had detailed mappings between the source object in the analogy and the relevant aspects of the science concept. This makes those studies atypical of how analogies are presented in textbooks.

The current paper presents two experiments that tested for the effects of analogies on JOUs about science concepts. In the first experiment, students judged their understanding of specific concepts, each presented in one or two sentences from a geoscience textbook with or without analogies. The second experiment tested the effects of those same analogies when they were embedded in longer texts about the science concepts and JOUs were made for the whole texts rather than individual sentences.

Experiment 1

Methods

Sixty-one participants (69% female, M_{age} = 18.7) who had not completed a college-level course in Geoscience participated in partial fulfillment of their course requirement for Introductory Psychology. In a between-participants design, for half of the participants the statements included a brief analogy (which for illustrative purposes is shown in italics in these examples). The other half saw no analogies.

All participants read 37 short, one or two sentence statements about specific geology concepts. The statements were edited versions of sentences found in a textbook widely used in Introductory Geoscience courses at the college level (Marshak, 2022). Most of the statements discussed causal processes in geology (e.g., "Magma has a lower density than solid rock, so it behaves buoyantly and rises, *like oil rises above vinegar in salad dressing*."). Some of statements were more descriptive about geological formations or events (e.g., "The Himalayas rise about 8km above sea level, and most of the thickened crust extends downward beneath the range, *just as most of a floating ice cube is under water*." Participants were told, "For each geology concept, you will rate how likely you are to understand a short text about it." They rated each concept on a 1 (Not at all) to 5 (Extremely high) scale of likelihood of understanding. Similar to analogies most commonly found in textbooks, the analogies were simple and did not include any explanation about how the everyday source objects mapped onto the geological concepts. Inclusion of the analogies made each statement 8.76 words longer on average and more grammatically complex, as indicated by ARI readability index (Analogy: M = 11.75, SD = 4.31; No Analogy: M = 10.11, SD = 4.09). There was no way to control for this added length without introducing other non-equivalences between the conditions. However, the lower readability predicts lower judgments, which is the opposite of the hypothesized increase.

Results

There was variance in average JOUs across the 37 individual geology concepts, ranging from 2.64 to 3.64 on the 5-point scale. Among the 666 bivariate coefficients, there was enormous variability ranging from .00 to .81, Mdn = .45, M = .48. There was also between-person variance in JOUs for each concept, with the lowest and highest possible ratings being observed for all 37 of the geology concepts. Despite this between-person and between-concept variance, and inconsistent and modest covariance, the analogy effect was robust as evidenced by the fact that there was not a single one of the 37 concepts where the average JOU was lower when presented with an analogy than without it.

The ARI readability scores for each statement (which reflect shorter length and lower grammatical complexity) were positively correlated with average JOUs for each statement, within both the Analogy and No Analogy conditions, rs(36) = .38 and .47, ps < .05. In contrast to this positive relationship, the statements that included analogies received higher JOUs despite being lower in readability.

To facilitate a more direct comparison to the results of Experiment 2, the 37 individual concepts were clustered by the 6 topics (e.g., Earthquakes, Glaciers) corresponding to the 6 different segments of textbook used in Experiment 2. This resulted in each participant having an average JOU score for each of the 6 topics. A repeated-measures ANOVA was conducted with analogies as the between-participant factor and topic as the repeated measure.

As seen in Figure 1, there were main effects for analogy condition, $F(1, 59) = 8.31$, $p < .01$, $\eta_p^2 = .12$, and for topic, $F(5, 295) = 29.52$, $p < .001$, $\eta_p^2 = .33$, but there was no interaction, $F(5, 295) = 1.20$, $p = .31$, $\eta_p^2 = .02$. Overall, the JOUs for geology concepts were significantly higher with analogies (M = 3.45, SD = .64) than without analogies (M = 2.96, SD = .78), and this pattern was consistent across the 6 topics with medium to large effect sizes (ds = .44 to .90). Despite the fact that average JOUs varied quite a lot for different topics, the presence of analogies consistently increased confidence on all topics. For the main effect of topic, simple contrasts revealed that Uplift received significantly lower JOUs than the 5 other topics, while Earthquakes received the highest JOUs, and Groundwater and Weathering were tied for the second highest JOUs, significantly higher than the other three topics. These patterns across topics are noted for later comparison in Experiment 2.

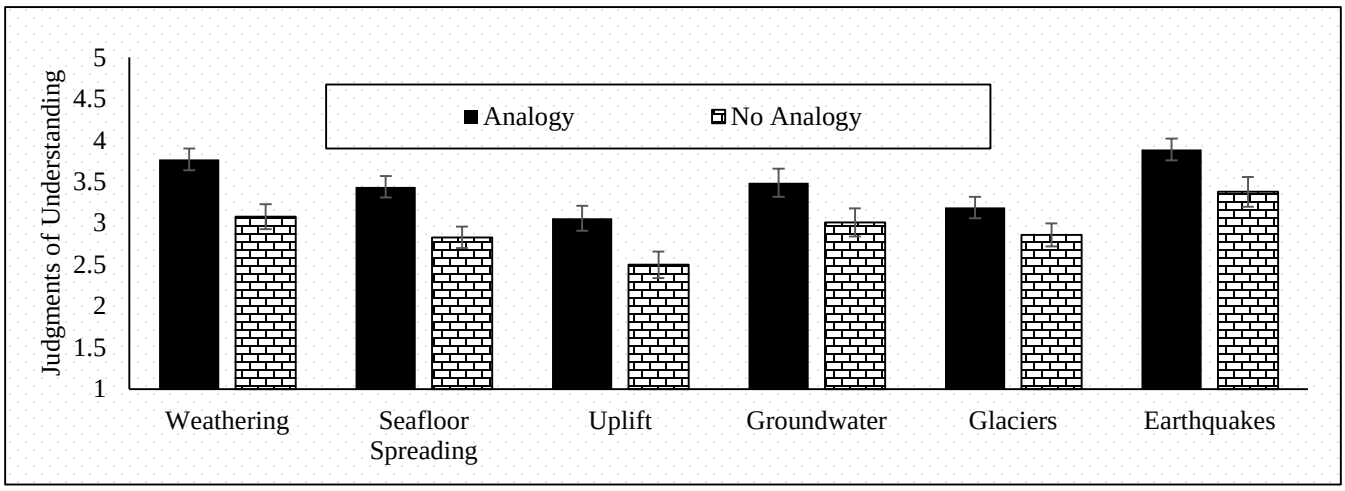

Figure 1: Experiment 1 JOUs for geology concepts with and without analogies

Experiment 2

The second experiment tested whether the same analogies would inflate more global JOUs for longer texts, each containing multiple analogies and explanations about the geology concepts. The same sentences taken from textbooks and used in Experiment 1 were now presented within the larger segment of the textbook in which they occurred. The texts were on the same 6 geology topics, each covering the target concepts used in the analogies plus closely related concepts. Importantly, while the longer texts included more explanation of the target geology concepts, they did not provide any additional elaboration on the mapping between the source objects and the concepts. Thus, the inclusion of analogies was superficial similar to what is common in geology textbooks. The use of analogies within longer texts allowed for testing inference-level comprehension of the concepts, and thus allowed for the computation of overconfidence of the JOUs.

Although two prior studies (Jaeger & Wiley, 2015; Wiley et al., 2018) found no effects from analogies on JOU magnitude, the present study differed in that it presented multiple analogies without providing clear mappings. Thus, there was no clear hypothesis for the effect of analogies. On one hand, the robust increase in JOUs observed in Experiment 1 from superficially presented analogies may persist even after reading a longer text about geological processes. On the other hand, reading a longer text may be sufficient to alert students to the fact that the familiarity of the source objects may not be a good indicator of their understanding of explanations about the geology concepts.

Methods

Fifty-four undergraduate students (68% female, M_{age} = 18.9) who had not completed a college-level course in Geoscience participated in partial fulfillment of a course requirement for Introductory Psychology. All participants read 6 texts about the same 6 topics shown in Figure 1. The same 37 statements used in Experiment 1 served as key sentences in the texts. Whether these sentences did or did not contain the superficial analogies was the sole difference between the texts for the two conditions. The texts were taken from subsections of the textbook, and edited to be between 900 and 1100 words in length. To increase the likelihood that readers would notice the analogies within these longer segments of text, each text both began and ended with one of the analogies, with additional analogies in between.

Participants made a JOU after reading each text by answering the prompt, "How many questions out of 5 are you likely to get correct on a test about [topic]". After reading and judging all 6 texts, participants completed a 5-question multiple choice inference test for each of the 6 topics.

Each text covered a topic including several key concepts. The 5-item tests for each topic were designed to provide assessment coverage across the concepts in each text, rather than provide multiple assessments of the same particular concept. The questions were designed to require conceptual understanding of the important causal relations, including those related to the geology concepts that were part of the analogies. However, none of the questions directly mentioned the analogies or the various source objects used in those analogies. The purpose was to test understanding of the geology concepts, not the analogies themselves (since that is the understanding that the textbook itself was written to impart). The judgment prompt was worded in terms of number of items predicted correct out of 5, which allowed for a direct comparison with test performance to evaluate overconfidence levels with and without analogies.

Results

Judgments of Understanding (JOUs) The length of the texts was negatively correlated with average JOU magnitude (r = -.33, p < .05). To test the effect of analogies on JOUs, a repeated-measures ANOVA was conducted with analogies as the between-participant factor and topic as the repeated measure. As seen in Figure 2 (first pair of bars for each topic), there was a main effect for topic, $F(5, 260) = 8.76$, p <.001, η_p^2 = .14, but not for analogy condition, $F(5, 52) = 0.00$, p = .96, η_p^2 = .00, and no interaction, $F(5, 260) = 1.65$, p = .15, η_p^2 = .03. Consistent with the lack of analogy effect, one can see that there is no consistent pattern, with JOUs being slightly higher for the analogy condition on half of the topics and slightly lower on half of the topics (none being significant simple effects).

The different average JOUs for each topic showed a similar pattern to those in Experiment 1. Participants were the most confident in their understanding about Weathering, Groundwater, and Earthquakes, and least confident in Uplift, which includes the difficult concept of isostatic equilibrium. Simple contrasts showed that Uplift received significantly lower JOUs than all 5 other topics. Weathering, Groundwater, and Earthquakes were all significantly higher than the other 3 topics.

Tests of Understanding Scores on the comprehension tests for the 6 topics are shown by analogy condition in Figure 2 (second pair of bars for each topic). A repeated measures ANOVA revealed that there was a main effect for topic, $F(5, 260) = 13.46$, p <.001, η_p^2 = .21, but no effect of analogy condition, $F(5, 52) = 1.65$, p = .20, η_p^2 = .03, and no interaction, $F(5, 260) = 0.96$, p = .45, η_p^2 = .02. Not only did the analogies fail to improve understanding of the geology concepts, but the F value above 1 is due to the fact that for two of the texts (Weathering and Uplift) the presence of superficial analogies led to *decreases* in understanding with medium effect sizes that were significant or marginal (ps = .05 and .06, ds = .56 and .52).

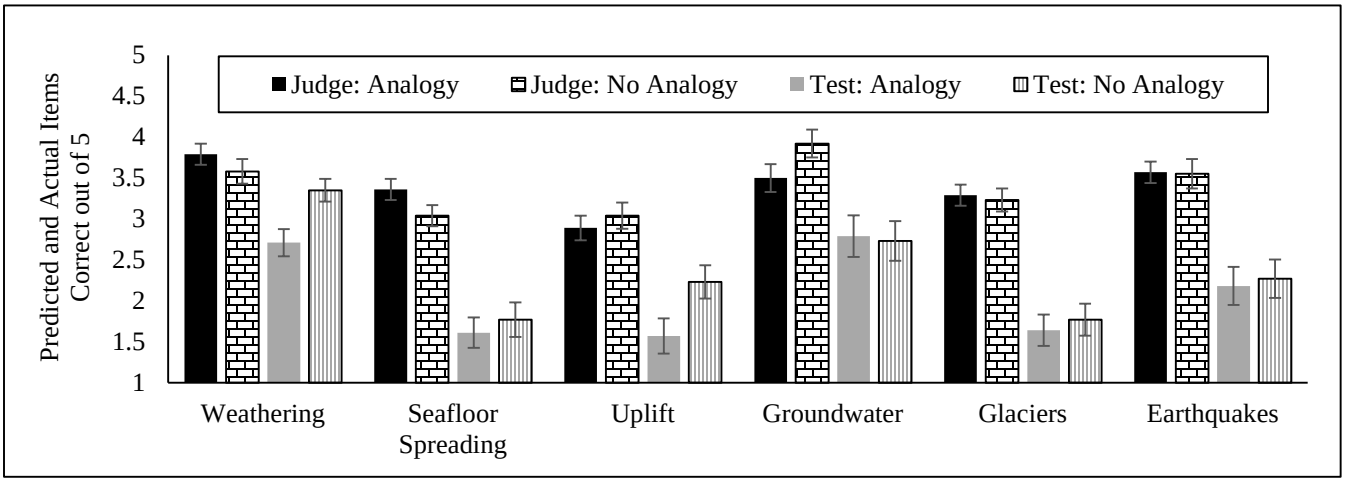

Figure 2: Experiment 2 judgments and test scores by analogy condition and topic

Follow-up tests revealed that differences in test scores between topics showed a similar pattern to the differences in JOUs. Weathering and Groundwater both had significantly higher test scores than all other topics, except for each other and Earthquakes. Uplift did not have the lowest test scores numerically, but it was not significantly higher than test scores on any other topic.

Overconfidence Figure 2 shows that the JOUs were notably higher on the 5-point scale than the actual test performance. On average, participants thought they would get 1.19 more questions correct out of 5 on each of the tests (SD = .95). This level of overconfidence was significantly greater than zero, $t(53) = 9.19$, $p < .001$. $d = 1.25$. Across all 30 test items, participants' overconfidence was equivalent to a difference of 24% on an exam.

Because there was no significant effect of analogy condition on either JOUs or test scores, there cannot be an overall effect on overconfidence. However, there is one topic where the effect of the analogy condition slightly increased JOUs while it significantly decreased test performance. For the Weathering topic, the degree of overconfidence (judgment-minus-test score) was significantly higher for the analogy condition (M = 1.07, SD = 1.44) than the no analogy condition (M = 0.23, SD = 1.24); $F(1, 52) = 5.24$, $p = .03$, d = .63.

Discussion

In the first experiment, the effect of higher JOUs for geology concepts that included analogies was strong and consistent across the various topics. Sentence length and complexity (as assessed by ARI) was negatively related to JOUs within each condition. However, something about the objects in the analogies was able to override that influence and lead readers to assume they had better understanding for the sentences with analogies, despite those being longer and more complex. The previously demonstrated (Griffin et al., 2024) greater familiarity, imageability, and concreteness of the everyday source objects used in geology analogies are likely reasons why their added length and grammatical complexity had the opposite effect than those variables more generally do. Although Experiment 1 did not measure actual understanding, the results from the comprehension tests in Experiment 2, including the overall poor test performance and higher overconfidence, suggest that the inflated JOUs in Experiment 1 are unlikely to reflect a similar degree of increase in real understanding.

The results of Experiment 2 suggest that this effect of analogies is not observed for more global JOUs made after reading longer texts that contain explanations of the geology concepts. These results were found in a context that might have been expected to result in an inflated sense of interest, familiarity or fluency, with multiple analogies mentioned only superficially (without a clear mapping between sources and targets). On the other hand, the average of 6 analogies per text added about 64 words. Since text length is negatively correlated with JOUs, it is possible that the compounded added length and complexity was enough to work against the sense of familiarity with the source objects. In addition, all texts, regardless of whether they contained analogies, had more in-depth explanations about the concepts plus additional related concepts. The statements judged in Experiment 1 comprised only about 25% of the sentences in the texts used in Experiment 2. These were naturalistic textbook segments with a typical ratio of analogized concepts. All the added information may overshadow the analogies when making global text-level judgments.

Experiment 2 extends the prior findings which also failed to see differences in JOUs from analogies in studies using texts with only a single analogy that was mapped out in detail in relation to the key concept (Jaeger & Wiley, 2015; Wiley et al., 2018). For one of the six topics, there was evidence of analogies leading to overconfidence, but similar to Jaeger and Wiley (2015), this was driven mostly by the negative effect of those analogies on actual understanding.

The overall consistently high level of overconfidence in Experiment 2 suggests that one aspect of the IOED studies may need to be reinterpreted. In some of those studies (e.g., Rozenblit & Keil, 2002), after people attempt to explain the everyday object, they then read an expert explanation and rate how well they understand the object now due to reading that expert explanation. People's final JOUs after reading rise back up to and sometimes even significantly higher than the initial JOUs prior to generating their own explanation. The authors of those studies did not interpret that final rise in JOUs as problematic, but merely as a "manipulation check", showing that people actually did read and understand the expert explanations, which is why they perceived their own earlier explanation as inferior by comparison. However, the current results suggest that students typically do not understand expert explanations as well as they believe they do. Therefore, even though generating explanations may help people realize that their initial confidence about everyday objects was an illusion, that illusion is easily revived after reading expert explanations, because people assume that they understand those explanations better than they actually do.

Wilson and Marsh (2023) recently found that a strong predictor of the drop in JOUs after explanation generation was people's perception that their own explanation was "incomplete". This was based on people's response to "What percent of the possible information do you think you produced?" (Wilson & Marsh, 2023, p. 719). Replies to that question are likely to reflect rather superficial qualities, such as the mere length of what they had generated. This is supported by the fact that subjective "completeness" ratings had only a weak correlation (r = .33) with an objective measure of completeness (actual number of parts and causal links assessed by coders). Likewise, readers may boost their JOUs after reading an expert explanation merely because that explanation is longer than their own, rather than because they actually understood the expert explanation. In sum, the post-explanation generation drop in JOUs may not reflect that the illusion of understanding has truly been dispelled or "broken", but merely superficially suspended in a way that allows it to be easily reinstated by shallow processing of an expert explanation of the object.

The current results also add to the evidence about how analogies impact actual learning. The lack of positive effects of analogies on understanding, plus the negative effect on understanding for two topics, is consistent with claims that analogies are unlikely to be beneficial without the learner being helped to map the elements of the source object onto the target concept (i.e., Gentner, 1983; Gilbert, 1989; Jee et al., 2010). This is important, given that widely used textbooks and online teaching resources often use analogies but fail to provide that kind of mapping most of the time (Poe et al., 2024). Efforts by textbook creators to make the science seem more interesting and relevant via analogizing it to things in everyday life that students are familiar with may sometimes undermine learning, despite any potential boost to student interest.

The fact that the central result in Experiment 2 is a null effect of analogies on JOUs could raise some concerns about the reliability and validity of the JOUs. However, those concerns can be addressed by several other findings. First, the effect of analogies on JOUs in Experiment 1 was highly consistent across the dozens of concepts, and these effects were medium to large in size for each of the 6 topics. Plus, both convergent and divergent validity of the JOUs were supported by the fact that JOUs were predicted as expected by readability indices in both conditions. In addition, between topic differences in JOUs should be influenced by stable factors such as some topics being generally more familiar to people and differences in the inherent conceptual complexities of the topics. Thus, the similar pattern for JOUs among topics across both experiments lends validity to the JOU measure. Finally, those consistent patterns of JOUs generally mapped onto the pattern of objective topic difficulty as indicated by the differences in test performance.

It should also be noted that the sizable inflation of judgments observed in Experiment 1 but not Experiment 2 does not necessarily mean that there will not be negative metacognitive effects of analogies when students are studying using their textbooks. Students may often merely skim their textbook, reading only key sentences, which is similar to the procedure in Experiment 1. They may only adopt an approach similar to Experiment 2, reading the full sections of their textbook, when they encounter something that seems unfamiliar. And the illusion of understanding created when skimming key sentences with analogies could cause them to miss when the science concept is unfamiliar and deserving of more in-depth processing.

Similarly, even when embedded within larger text segments, the analogies could have a metacognitive impact that impedes understanding. The analogies may create an initial illusion of understanding that, despite not manifesting in post-reading global JOUs about the entire text, could persist long enough during reading to alter text processing in ways that impair understanding. Self-regulated learning behaviors during reading are impacted by attention to meta-experiences that are generated during text processing, and prior familiarity is a heuristic cue that can undermine attention to such meta-experiences (for a review, see Griffin, Wiley, & Salas, 2013). Also, metacognitive effects on self-regulated learning are likely to be stronger in situations where learners have the opportunity and motive to select texts for multiple restudy sessions. Future research needs to examine these issues using behavioral measures of text processing, including eye tracking and restudy. In addition, it will be important to explore whether the inflated judgments due to analogies observed in Experiment 1 can be lessened with instructions that both teach students about the nature of analogies and their unreliable impact on understanding, depending on how they are presented and the extent to which they are explained.

Acknowledgements

The authors thank Allison Jaeger and Nicole LeDue who are Co-PIs on the larger collaborative project supported by the National Science Foundation under Grant EHR 2300990. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Coltheart, M. (1981). The MRC psycholinguistic database, *Quarterly Journal of Experimental Psychology, 33A*, 497-505.

Curtis, R.V., & Reigeluth, C.M. (1984). The use of analogies in written text. *Instructional Science, 13,* 99-117.

Donnelly C. M., & McDaniel M. A. (1993). Use of analogy in learning scientific concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 975-987.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271-280.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 552-564.

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

Gilbert, S. (1989). An evaluation of the use of analogy, simile, and metaphor in science texts. *Journal of Research in Science Teaching, 26*, 315-327.

Griffin, T. D., Guerrero, T. A., Mielicki, M., & Wiley, J. (2025). Improving self-regulated learning of less-prepared college students with lessons about inferences. *Journal of Applied Research in Memory and Cognition.* Advance online publication. https://doi.org/10.1037/mac0000207

Griffin, T. D., Jaeger, A., LaDue, N., Adams, R.L., & Wiley, J. (2024, July). *Analogies in geology textbooks may create illusions of understanding.* Poster presented at the 34th Annual Meeting of the Society for Text and Discourse, Chicago, IL.

Griffin, T. D., Mielicki, M. K., & Wiley, J. (2019). Improving students' metacomprehension accuracy. In J. Dunlosky & K. Rawson (Eds.), *Cambridge handbook of cognition and education* (pp. 619-646). Cambridge University Press.

Griffin, T. D., Wiley, J., & Salas, C. (2013). Supporting effective self-regulated learning: The critical role of monitoring. In R. Azevedo & V. Aleven (Eds.) *International handbook of metacognition and learning technologies* (pp. 19-34). Springer Science.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2019). The effects of comprehension-test expectancies on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*, 1066-1092.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*, 160-170.

Hildenbrand, L., Sarmento, D., Griffin, T. D., & Wiley, J. (2024). Conceptual overlap among texts impedes comprehension monitoring. *Psychonomic Bulletin & Review*, 31, 750-760.

Jaeger, A. J. & Wiley, J. (2015). Reading an analogy can cause the illusion of comprehension. *Discourse Processes, 52*, 376-405.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T., Sageman, B., Ormand, C. J., & Tikoff, B. (2010). Analogical thinking in geoscience education. *Journal of Geoscience Education, 58*, 2-13.

Marshak, S. (2022*). Earth: Portrait of a planet (7th Edition)*. W. W. Norton & Company.

Orgill, M. & Bodner, G.M. (2006), An analysis of the effectiveness of analogy use in college-level biochemistry textbooks. *Journal of Research in Science Teaching, 43*, 1040-1060.

Poe, M. D., Jaeger, A. J., Sandefur, L., & LaDue, N. (2024, December). *The Earth is like peach: Understanding the presence of analogical examples in introductory geoscience texts*. Poster presented at the Annual Meeting of the American Geophysical Union. Washington, D.C.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*, 521-562.

Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*, 698-711.

Thiele, R.B., & Treagust, D.F. (1994). The nature and extent of analogies in secondary chemistry textbooks. *Instructional Science, 22*, 61-74.

Wiley, J. (2019). Picture this! Effects of photographs, diagrams, animations, and sketching on learning and beliefs about learning from a geoscience text. *Applied Cognitive Psychology*, *33*(1), 9-19.

Wiley, J., Jaeger, A. J., Taylor, A. R., & Griffin, T. D. (2018). When analogies harm: The effects of analogies on metacomprehension. *Learning and Instruction, 55*, 113-123.

Wilson, J., & Marsh, J. K. (2023). Perceptions of explanation completeness help decrease knowledge overestimation. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 45*, 717-723.