Explaining "Why" Matters: Mechanistic Explanations of Misconceptions Promote Learning of Counterintuitive Conceptual Content

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Elgamal, Karima
Brown, Sarah A
Kelemen, Deborah

Publication Date

2026

Explaining “Why” Matters: Mechanistic Explanations of Misconceptions Promote Learning of Counterintuitive Conceptual Content

Karima Elgamal (karima02@bu.edu), Sarah Brown (sarah.brown027@gmail.com), Deborah Kelemen (dkelemen@bu.edu)
Department of Psychological & Brain Sciences, Boston University

Abstract

What type of information best supports learning? In educational settings, instructional information may describe patterns of error and provide corrections, or explain the mechanisms that generate those errors. This distinction may be especially consequential in scientific domains where misconceptions are common. We tested whether mechanistic explanations of why misconceptions arise support adults’ learning of a counterintuitive scientific idea beyond descriptive error correction. U.S. elementary school teachers (N = 127) were assigned to a Control, Description, or Mechanistic explanation condition in a professional development intervention on natural selection, where they heard different accounts of common student misconceptions. Teachers’ generalized learning of natural selection was assessed using near and far transfer measures. Near generalization improved across all conditions, but far generalization improved only in the Mechanistic condition, suggesting that fostering a deeper conceptual framework for understanding misconceptions supports broader, more flexible learning of counterintuitive conceptual content than descriptive correction.

Keywords: learning; explanation; mechanistic explanation; misconceptions; intuitive bias; natural selection; science education

Introduction

Across development, learners encounter substantial amounts of explicit instructional information intended to correct errors that might undermine learning. However, this information varies not only in form but potentially also in how it functions - particularly in domains involving counterintuitive concepts, where learners often hold persistent misconceptions. We explored whether the form of misconception correction shapes learning outcomes in a teacher professional development intervention that used the counterintuitive idea of natural selection as a test case. To address this question, we contrasted two forms of instruction: descriptive information and explanatory information.

Descriptive information characterizes observable outcomes and patterns, or simply states “what is correct.” In science education, this approach commonly appears in misconception-focused instruction, such as refutation text, which explicitly presents a misconception, labels it as incorrect, and provides corrective information intended to replace it with the scientific view (Tippett, 2010). Refutation-style instruction can support learning and generalization (Schroeder & Kucera, 2022), potentially because targeted focus on correct versus incorrect ideas alone is efficient at prompting a cognitive conflict that encourages revision or suppression of previously endorsed inaccurate beliefs. By juxtaposing an incorrect claim with a corrective statement, refutation texts make the contrast between competing ideas salient, efficiently supporting the kind of pattern extraction that may enhance recognition of, and monitoring for, specific erroneous propositions. It is a form of instruction that blurs the distinction between description and explanation. However, while identifying a particular idea as incorrect and stating why it fails in a specific case can provide a localized explanation of error, it still falls short of clarifying the underlying causal processes that generate the broader error pattern (Horwood, 1988).

The distinction becomes clearer when considering mechanistic explanations, which account for how and why a phenomenon occurs by articulating the interacting processes that generate it (Glennan, 2002; Lombrozo, 2010). By making underlying causal structure explicit, highlighting which features are causally relevant as well as how they emerge and interact, mechanistic explanations may provide learners with a framework for organizing information, evaluating competing accounts, and generating predictions beyond the instructional context. However, mechanistic explanations of conceptual error may not always facilitate learning: relative to targeted corrective descriptions, they can be longer and more complex, offering details that learners may experience as extraneous or hard to integrate. Thus, causal-mechanistic explanations may either support or hinder broader generalized learning due to their time and cognitive processing demands.

Misconceptions

A consideration of the underlying nature of scientific misconceptions is also relevant to predictions about particularly effective ways to correct them (Chi, 2013).

Under one account, most misconceptions arise because knowledge is fragmented: errors occur because learners lack a relevant fact or misremember a specific relationship (e.g. Bhatia et al., 2022; diSessa, 2018; Scherr, 2007).

Under another account, misconceptions are often not random slips or isolated factual errors, but more systematic byproducts of intuitive cognitive tendencies. From early in development, people construct theory-like frameworks to make sense of the world, and although these intuitive frameworks are often adaptive for everyday reasoning, they can lead to systematic misinterpretations when learning scientific ideas that appeal to conflicting abstract

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

frameworks of assumptions (Carey, 1985; Vosniadou, 1994; Kelemen, 1999; Shtulman & Harrington, 2016). Because misconceptions may have a degree of intuitive theoretical coherence, they can be hard to dislodge or over-ride. Thus, even after instruction, they may reemerge in unfamiliar contexts or under cognitive load (Kelemen, 2019; Kelemen & Rosset, 2009).

These two characterizations of misconceptions generate somewhat competing predictions about instruction and learning. If misconceptions frequently reflect relatively incoherent fragments of knowledge, then descriptive error correction - identifying incorrect ideas and providing accurate alternatives - may be sufficient for robust learning. This may also be because the targeted content is cognitively efficient: when an isolated idea comes to mind, learners may easily pattern match to a corrective refutation to reject it.

However, if misconceptions reflect systematic intuitive biases that yield theory-like frameworks, then learners may benefit from instruction that makes both the causal structure of a target concept and the origins of intuitive error explicit. This is because, while they may contain apparently extraneous information that takes more effort to process, explanations can serve as organizing elements in conceptual representations (Carey, 1985; Murphy & Medin, 1985; Lombrozo, 2009). Mechanistic explanations may be especially well suited to supporting generalized learning by (1) restructuring how learners interpret new evidence and (2) by making the origins of intuitive errors intelligible. The latter could possibly support a metacognitive awareness of misconception-producing tendencies, facilitating inhibition of those defaults in novel contexts. Although prior work suggests that instruction emphasizing underlying structure can promote transfer in some contexts (Lombrozo & Carey, 2006; Rittle-Johnson, 2006), it is unclear whether explicitly explaining the cognitive origins of misconceptions provides benefits beyond descriptive correction, in domains characterized by persistent intuitive biases.

Test Case: Natural Selection

Biological evolution by natural selection provides an especially powerful test case for examining the most effective instructional approach to error correction. Misconceptions about evolution are pervasive across age groups, cultures, and levels of expertise (Nehm & Schonfeld, 2008; Gregory, 2009; Kelemen & Rosset, 2009; Kelemen, et al., 2013). A common error involves reasoning about biological traits in teleological terms such as explaining biological traits in terms of purpose or need (e.g., “Polar bears evolved white fur so they could camouflage in the snow”). Such ideas are highly intuitive but fundamentally incompatible with the non-goal-directed mechanisms of natural selection (Kelemen, 1999; Shtulman & Calabi, 2012).

Instruction on natural selection typically combines accurate mechanistic accounts of evolutionary processes, with efforts to reduce teleological errors through descriptive identification and correction of incorrect ideas. If the misconceptions simply reflect relatively fragmented or locally misapplied ideas, such targeted corrective descriptions may be sufficient for supporting accurate generalizable learning. However, if these errors reflect systematic explanatory patterns that extend across contexts, correcting individual responses without an explanation for why this type of error arises may not support accurate and generalizable reasoning. Natural selection therefore provides a domain for testing whether learning benefits from instruction that communicates the correct causal structure and explains why certain patterns of error arise.

Current Study

The present study tested the benefit of different kinds of misconception instruction in a consequential group of adult learners: elementary school teachers. This is a population of non-science specialists whose understanding of evolution directly influences how young students first encounter the foundational idea of natural selection and whether they begin to entrench inaccurate ideas about it. Teachers commonly encounter intuitive misconceptions in student reasoning (Nehm & Schonfeld, 2008), making it critical that they can generalize a scientifically accurate understanding. They were therefore an ideal sample for studying, whether supplementing instruction on natural selection with a mechanistic (developmental) explanation of the origins of misconception improves generalized learning beyond descriptive error correction alone.

Method

Participants

The final sample was 127 elementary school teachers recruited via email from public school district lists across the United States. Nine participants were excluded for failing attention checks or incomplete participation.

Participants were mostly female (91%; M age = 37 years, SD = 12, range = 22-68). Most identified as White (83%), then Black or African American (13%), Asian (2%), multi-racial (2%), <1% as Native American or Alaskan Native; 6% identified as Hispanic or Latino. Participants represented 15 states, with the largest proportion in Massachusetts (33%). This distribution closely mirrors national demographics of U.S. elementary school teachers, who are approximately 80% female and 80% White (National Center for Education Statistics, 2023).

Design & Procedure

The study employed a pretest-intervention-posttest design with random assignment to one of three between-subjects conditions: Control, Description, or Mechanistic Explanation (OSF: https://osf.io/gzqu4/overview). All components were delivered asynchronously via Qualtrics links as a self-paced professional development course and completed over two to three days (median = 2), with 90% completing all components within one week. This method is adapted from an extended professional development intervention on explanations from Elgamal, et al. (*in preparation*).

Core Instructional Materials In all conditions, the intervention consisted of three core instructional segments. First, teachers viewed a 4-min long introductory video on the importance of teaching natural selection, which also noted that students often misunderstand it. In a second 8-min video, teachers were then taught an accurate population-based mechanistic account of adaptation by natural selection via a narrated animated version of the children's storybook *How the Piloses Evolved Skinny Noses* (Kelemen et al., 2014). The story explains how variation in nose shape in a novel mammalian species led to differential access to food after an environmental change. Differential survival and reproduction across generations then prompted change in the species' predominant nose shape over time. Later, all teachers also viewed a final 3-min recap video that summarized the key causal elements of the storybook and explained the instructional rationale underlying its design.

Condition Specific Intervention The experimental conditions differed only in a single condition-specific video segment that appeared between the storybook reading and the final recap video. This segment was designed to manipulate the form of information teachers received about students' misconceptions after all teachers received the same scientific explanation of natural selection. This sequence therefore allowed us to isolate the effect of the way misconceptions were presented to teachers on their near and far generalization of natural selection.

In the Control condition, teachers viewed a brief (~3 min) video presenting accurate information about natural selection and general guidance about accurately teaching the topic. The materials emphasized core components of natural selection (e.g., variation) and noted that students often hold incorrect ideas about the process that teachers should identify and correct. Misconceptions were therefore framed simply as an important instructional challenge. The information did not include systematic instruction about specific misconception patterns or analysis of the reasoning that produces them.

In the Description condition, teachers viewed a longer (~10 min) video that included the same core scientific instruction as the Control condition, followed by a series of examples of student misconceptions centered on purpose-based (teleological) reasoning. Teachers were shown student explanations that appealed to need or goal-directed change (e.g., "the animals changed so that they could survive"), and each was refuted through a description of why this reasoning is scientifically inaccurate.

In the Mechanistic Explanation condition, teachers viewed the same core scientific instruction and the same misconception examples and corrective explanations as in the Description condition (~10 min). However, prior to the misconception examples, this condition included an additional video (~7 min) that provided a mechanistic account of why these misconceptions are common. Specifically, teachers were introduced to the cognitive developmental idea that, from early in development, people construct intuitive explanatory tendencies that help them make sense of the world, and that one of these - a purpose bias - can lead learners, from early on, to broadly favor goal-directed explanations of many kinds of phenomena including biological change. This segment emphasized that such explanations often feel compelling and cognitively "natural," framing teleological errors not simply as incorrect ideas but as predictable products of intuitive reasoning. Teachers were encouraged to use this framework to recognize when an explanation's surface plausibility might stem from intuitive bias rather than scientific validity. This bias was referenced again as example errors were later described and corrected, thereby reinvoking the mechanistic explanation of their origins.

Assessments (Pre vs. Post) At pre- and post-test, teachers completed two measures of conceptual understanding of natural selection: the Evolving Minds (EM) Adaptation Assessment (Kelemen et al., 2014) (which was designed for children and served as a measure of near generalization at post-test due to its structural alignment with the storybook), and the multiple-choice Conceptual Inventory of Natural Selection (CINS; Anderson et al., 2002) (which, at post-test, served as a measure of far generalization because its format and scenarios differ substantially from the storybook).

The EM assessment presented a novel species before and after an environmental change, showing a shift in the frequency of a particular trait (e.g., from mostly shorter arms to mostly longer arms). Participants ability to use natural selection to explain the trait change was then measured through a combination of six forced-choice questions, plus justifications that targeted isolated components of natural selection (e.g., differential survival, reproduction, inheritance, and trait constancy), (see Table 1 for sample items), and three prompts to offer an open-ended explanation of the change (see Figure 1 for open-ended pretest questions).

The CINS (Anderson et al., 2002) is a scenario-based 15 multiple-choice measure assessing conceptual understanding of evolution by natural selection (8 questions on finches and 7 questions on lizards; Table 2).

Immediately following the intervention, an EM assessment and CINS posttest were re-administered.

Data Coding

EM Adaptation Assessment Responses on the EM assessment were coded using a manual adapted from Kelemen, et al. (2014). Participants were dichotomously classified as demonstrating a population-based understanding of natural selection if their responses met three criteria: (1) correct answers with accurate justifications on at least five of six forced-choice items; (2) absence of misconceptions across the assessment, including teleological claims (e.g., change due to need or purpose); (3) open-ended explanations describing adaptation via differential survival at minimum. This coding had high inter-coder reliability ($\kappa \geq .80$) and yielded two binary near generalization learning outcomes for analysis: presence or absence of a natural selection misconception; presence or absence of an accurate population-based explanation of natural selection.

CINS Correct responses to the fifteen CINS multiple choice questions were summed to create a total score (0-15 range). Items left blank were coded as incorrect.

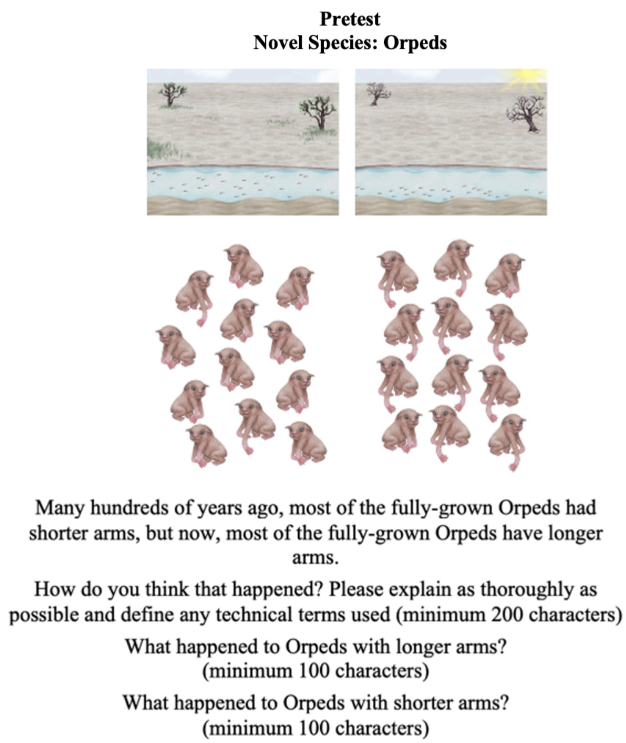

Figure 1. Open-ended questions for the pretest EM Adaptation Assessment.

Pretest	Posttest
Will an Orped with shorter arms probably be healthy and live for a long time? [Yes/No] *Why?*	Will a Tardon with a stumpier tail probably be healthy and live for a long time? [Yes/No] *Why?*
Will an Orped with longer arms probably have lots of children? [Yes/No] *Why?*	Will a Tardon with a stretchier tail probably have lots of children? [Yes/No] *Why?*

Table 1. Sample forced choice and justifications items from the EM Adaptation Assessment for pre and posttest.

Scenario	Question	Response Options
Finch Scenario Example	What would happen if a breeding pair of finches was placed on an island under ideal conditions with no predators and unlimited food so that all individuals survived? Given enough time…	a) The finch population would stay small because birds only have enough babies to replace themselves. b) The finch population would double and then stay relatively stable. c) The finch population would increase dramatically. d) The finch population would grow slowly and then level off.
Lizard Scenario Example	What could cause one species to change into three species over time?	a) Lizards in each group learned new behaviors. b) Lizards in each group adapted to different environments over generations. c) Lizards in each group changed to match the food available. d) Lizards in each group decided to become different species.

Table 2. Subset of items from the Conceptual Inventory of Natural Selection items (Anderson et al., 2002).

Results

EM Adaptation Assessment

At pretest, more than half of teachers (59%) did not understand natural selection such that they actively expressed a misconception This misunderstanding dropped to 24% at posttest, a significant reduction, McNemar's $\chi^2(1, N = 127) = 37.96$, $p < .001$. However, logistic regression indicated that while pretest performance predicted posttest misconception status ($b = 0.30$, $p < .001$), there was no effect of condition and no significant pairwise differences, $\chi^2(2) = 0.25$, $p = .88$.

Teachers' production of accurate population-based explanations of natural selection showed a parallel pattern. At pretest, only 32% of teachers produced a population-based explanation. This increased to 71% by posttest, $\chi^2(1, N = 127) = 40.42$, $p < .001$. As in the misconception analysis, condition did not predict posttest performance when controlling for baseline understanding, $\chi^2(2) = 1.55$, $p = .46$, and no pairwise comparisons were significant. Thus, all three intervention conditions equivalently supported near generalization from the instructional materials to the structurally aligned assessment, suggesting that exposure to the storybook might have been all that was needed.

CINS

Pretest CINS scores reflected teachers' limited baseline understanding of natural selection and did not differ across conditions (Control (M = 8.02, SD = 3.12), Description (M = 7.33, SD = 2.76), and Mechanistic (M = 7.86, SD = 3.04)), F(2, 124) = 0.62, p = .54. These CINS scores increased from pretest to posttest, t(126) = 7.24, p < .001, confirming learning across all conditions(Control, M = 9.14, SD = 3.47; Description, M = 9.52, SD = 3.26; Mechanistic, M = 10.71, SD = 3.49; Figure 2).

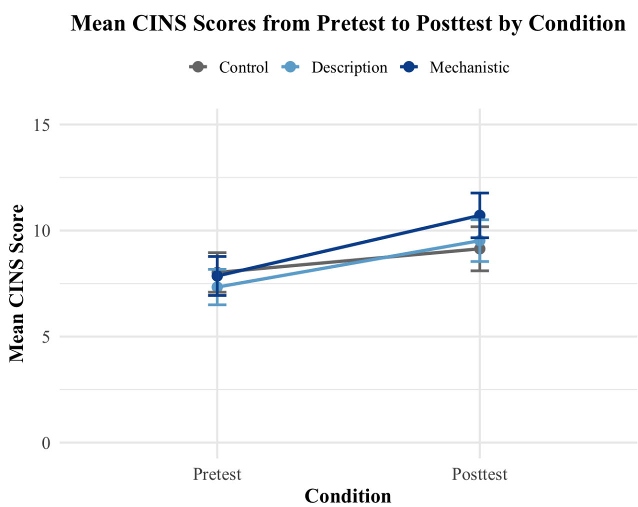

Figure 2. Mean CINS scores at pretest and posttest by condition. Error bars show ±1 SD. Teachers improved from pretest to posttest across conditions, with no baseline differences and highest posttest performance in the Mechanistic condition.

To test the effect of instructional condition on posttest performance, we regressed posttest CINS scores on condition while controlling for pretest performance. The model was significant, $F(3, 123) = 18.02$, $p < .001$. Pretest CINS score was a strong predictor of posttest performance (b = 0.60, p < .001; see Figure 3). Controlling for pretest, teachers in the Mechanistic Explanation condition scored higher than those in the Control condition (b = 1.67, p = .009), whereas the Description condition did not differ from Control (b = 0.80, p = .209). The Mechanistic and Description conditions also did not significantly differ from one another (b = 0.87, p = .14).

In sum, teachers in the Mechanistic Explanation condition outperformed those in the Control condition on far generalization. Performance in the Description condition did not significantly differ from either the Control or Mechanistic conditions.

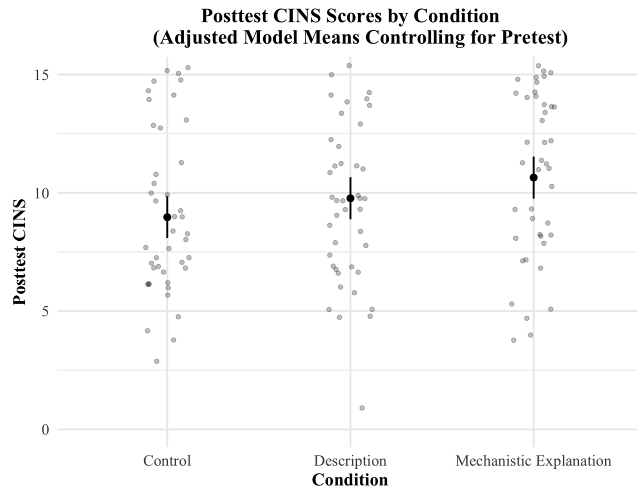

Figure 3. Points and error bars represent estimated marginal means and 95% confidence intervals from a linear model predicting posttest CINS scores from condition while controlling for pretest performance. Faint points show individual raw posttest scores.

Discussion

The present study tested whether the functional form of instructional information shapes learning outcomes in a teacher's professional development intervention on natural selection. All participants received the same accurate instruction on natural selection, but conditions varied in whether additional content provided descriptive misconception error correction (Description) or mechanistic accounts of why misconceptions arise (Mechanistic Explanation). This manipulation was motivated by competing accounts of misconceptions described in the introduction: if misconceptions primarily reflect isolated gaps or fragments of knowledge, then correcting incorrect ideas and supplying accurate alternatives should be sufficient for robust learning. If, however, misconceptions reflect systematic products of intuitive cognitive tendencies, then instruction may need to do more than correct responses, specifically, learners' may need a theory to compete with and suppress default theory-based misconceptions about counterintuitive ideas, such that they engage in accurate and flexible reasoning after instruction.

Divergent Patterns

Performance on the EM Adaptation Assessment improved comparably across conditions. This near generalization measure was closely aligned with the intervention's representational format and causal structure, allowing teachers to succeed by applying recently taught principles within familiar task constraints. In this context, improvement appears to reflect successful retrieval and application of the storybook's mechanistic account of natural selection itself, suggesting that the child-appropriate mechanistic explanation alone was sufficient to support near generalization, regardless of how misconceptions were subsequently addressed across conditions.

A different pattern emerged on the CINS. Unlike the structurally aligned EM adaptation measure, the CINS involved different evolutionary scenarios and the need to discriminate multiple choice response options provided weaker surface cues linking back to instruction and thus a far generalization test. Only teachers in the Mechanistic Explanation condition showed greater learning gains relative to the control condition, suggesting that the added value of explaining the origin of misconceptions emerges most clearly when learners must recruit natural selection principles far beyond the trained format.

Why The Different Patterns?

This dissociation can be understood in terms of two functional differences between conditions. First, they differed in how misconceptions were conceptualized: the Descriptive condition treated them as isolated incorrect ideas to be replaced, whereas the Mechanistic condition framed them as interconnected products of underlying cognitive tendencies that make certain explanations feel compelling. Second, the level of abstraction differed: descriptive correction operated at the level of specific items, whereas mechanistic explanation addressed domain-general explanatory tendencies, offering a framework for reasoning about explanation itself that could transfer beyond the instructional examples. Together, these differences suggest that mechanistic explanation of misconception origins functions not merely as additional content, but as a higher-order framework that becomes particularly useful when learners must select among competing explanations in unfamiliar contexts.

Possible Interpretations

One interpretation is that mechanistic explanations of misconception origins function as training on the *source* of errors, rather than merely their presence. By making explicit how intuitive reasoning tendencies give rise to particular explanations, this instruction may help learners recognize the underlying structure of distractor responses. In this way, the Mechanistic condition may have provided a transferable metacognitive framework, namely, that purpose-based reasoning can feel explanatory yet mislead in evolutionary contexts, supporting reorganization of knowledge structures and more flexible evaluation of evidence. Such a framework would be especially useful on tasks like the CINS, where surface cues are weak and explanatory competition is high.

A simpler alternative is that the mechanistic condition benefited from increased time-on-task. However, this account does not align well with the observed pattern of results. If additional exposure were the primary driver, improvements should have appeared across measures, including the structurally aligned EM near-generalization task. Instead, the advantage was selective to the far generalization CINS measure. Moreover, the added content in the Mechanistic condition did not further elaborate natural selection principles, but instead introduced abstract accounts of intuitive reasoning that were not directly tied to the target content. If anything, such material might have been expected to dilute learning rather than enhance it.

Taken together, these findings suggest that instructional effectiveness depends not only on *what* information is presented, but on whether it supports learners in regulating their own explanatory tendencies. While descriptive correction may suffice when tasks involve selecting among familiar causal models, explaining why misconceptions arise may provide a higher-order scaffold for reasoning in contexts where surface alignment is weak. In domains characterized by persistent intuitive biases, this type of instruction may be particularly important for supporting robust generalization.

Limitations & Future Research

If mechanistic explanations of misconception origins promote generalization by improving learners' monitoring of intuitive error, their benefits should be especially evident over time and in instructional contexts that require active regulation of intuitive thinking (which was not explored in this study). Exploring this prediction further would require delayed posttests and measures of instructional transfer, for example, testing whether effects persist months after the intervention and whether teachers become better at identifying and responding to students' teleological or intuitive explanations in classroom-like scenarios.

Future work is also needed to directly examine the proposed metacognitive mechanism. The present measures do not directly reveal the processes underlying the CINS advantage and did not measure teachers' awareness of intuitive error. Establishing that mechanistic accounts of misconception origins alter how learners monitor and regulate their own reasoning remains an important next step.

More broadly, these findings reinforce the paper's central claim: not all instructional information functions in the same way. Descriptive correction, and even simply hearing a correct explanation (i.e. the storybook) may be sufficient when tasks are closely aligned with instruction, but explaining the causal origins of misconceptions may be critical for transfer to new contexts. Lastly, this implication likely extends beyond evolution to other domains shaped by persistent intuitive defaults, such as naive physics (McCloskey, et al., 1983) or probability (Gilovich, et al., 1985), where learners may know the correct principle yet revert to intuitive explanations when contexts shift or cognitive demands increase. In such cases, instruction that explains why common errors arise may serve as a metacognitive scaffold, helping learners recognize and regulate intuitive reasoning and thereby supporting flexible conceptual generalization.

Acknowledgements

This research was funded by National Science Foundation DRL-2009176 and DRL-1561401 awarded to Deb Kelemen.

References

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. Journal of Research in Science Teaching, 39(10), 952–978.

Bhatia, K. S., Stack, A., Sensibaugh, C. A., & Lemons, P. P. (2022). Putting the Pieces Together: Student Thinking about Transformations of Energy and Matter. *CBE life sciences education*, *21*(4), ar60.

Baillargeon, R. (2004). Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental Science, 7*(4), 391–414.

Carey, S. (1985). *Conceptual change in childhood*. MIT Press.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*(2–3), 105–225.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*(S3), S342–S353.

Gregory, T. R. (2009). Understanding natural selection: Essential concepts and common misconceptions. *Evolution: Education and Outreach, 2*(2), 156–175.

Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.

Horwood, R. H. (1988). Explanation and description in science teaching. *Science Education, 72*(1), 41–49.

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227–254.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition, 70*(3), 241–272.

Kelemen, D. (2019). The magic of mechanism: Explanation-based instruction on counterintuitive concepts in early childhood. *Perspectives on Psychological Science, 14*(4), 510-522.

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition, 111*(1), 138–143.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General, 142*(4), 1074–1083.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences, 10*(10), 464–470.

Lombrozo, T. (2009). Explanation and categorization: How "why?" informs "what?". *Cognition, 110*(2), 248–258.

Lombrozo, T. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology, 61*(4), 303–332.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Mind & Language, 21*(5), 539–555.

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 636–649.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*(3), 289–316.

National Center for Education Statistics. (2023). *Characteristics of public school teachers*. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved May 6, 2026, from https://nces.ed.gov/programs/coe/indicator/clr

Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching, 45*(10), 1131–1160.

Quine, W. V. O., & Ullian, J. S. (1970). *The web of belief.* Random House.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*(1), 1–15.

Scherr, R. (2007). Modeling student thinking: An example from special relativity. *Am. J. Phys,* 75 (3), 272–280. https://doi.org/10.1119/1.2410013

Schroeder, N. L., & Kucera, A. C. (2022). Refutation text facilitates learning: A meta-analysis of between-subjects experiments. *Educational Psychology Review, 34*, 1109–1141.

Shtulman, A., & Calabi, P. (2012). Cognitive constraints on the understanding and acceptance of evolution. In K. S. Rosengren, S. Brem, E. M. Evans, & G. M. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 47–65). Oxford University Press.

Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*(1), 118–137.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science, 348*(6230), 91–94.

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8*, 951–970.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45–69.

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science, 34*(5), 776–806.