JOURNAL OF RESEARCH IN SCIENCE TEACHING

Research Article

Young Children's Near and Far Transfer of the Basic Theory of Natural Selection: An Analogical Storybook Intervention

Natalie Emmons,[1] Kristin Lees,[2] and Deborah Kelemen[1]

[1]*Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts*
[2]*Department of Population Health, Northeastern University, Boston, Massachusetts*

Received 6 March 2017; Accepted 10 August 2017

Abstract: Misconceptions about adaptation by natural selection are widespread among adults and likely stem, in part, from cognitive biases and intuitive theories observable in early childhood. Current educational guidelines that recommend delaying comprehensive instruction on the topic of adaptation until adolescence therefore raise concerns because children's scientifically inaccurate theories about species may be left unchallenged for many years, allowing them to entrench and become difficult to overcome. In consequence, this investigation sought to explore whether classrooms of kindergartners and second graders could acquire a basic but comprehensive understanding of adaptation from an intervention constructed around two picture storybooks that mechanistically explain natural selection. Learning was assessed in near and far transfer contexts both immediately and a month later. Kindergartners and second graders demonstrated substantial learning of biological information; however, second graders showed pronounced abilities to near and far generalize, immediately and over time. Results suggest that causally cohesive interventions with an emphasis on mechanistic explanation facilitate children's classroom learning of complex counterintuitive scientific ideas. © 2017 Wiley Periodicals, Inc. J Res Sci Teach 9999:XX–XX, 2017.

Keywords: evolution; natural selection; science education; cognitive development; cognitive biases; classroom intervention; analogy

Adaptation by natural selection is a fundamental and unifying concept in biological science. It helps to explain the existence and nature of all living things from species' physical traits to their behaviors, psychology, and even cultural practices. However, despite its theoretical importance and contributions to numerous fields including medicine, biotechnology, and environmental science, most people misunderstand how adaptation by natural selection occurs (see Gregory, 2009, for a review). Furthermore, misconceptions about the process are not easily remedied: They robustly persist following explicit and often lengthy instruction (Ferrari & Chi, 1998; Jensen & Finley, 1995; Vlaardingerbroek & Roederer, 1997) and often extend to teachers tasked with educating students on the subject (Asghar, Wiles, & Alters, 2007; Nehm & Schonfeld, 2007; Rutledge & Warden, 2000). The sources of these misunderstandings are numerous (see Rosengren, Brem, Evans, & Sinatra, 2012, for discussions); however, the role of everyday

---

Contract grant sponsor: National Science Foundation; Contract grant numbers: 1007984, 50204432.

Correspondence to: N. Emmons; E-mail: nemmons@bu.edu

DOI 10.1002/tea.21421

Published online in Wiley Online Library (wileyonlinelibrary.com).

© 2017 Wiley Periodicals, Inc.

"intuitive" reasoning biases observable from at least preschool age has gained increasing attention.

These early emerging, untutored cognitive biases, which include tendencies to essentialize and teleologically explain animal properties, guide our category-based judgments about animals (e.g., Coley & Tanner, 2012, 2015; Evans, 2000, 2001; Gelman & Rhodes, 2012; Kelemen, 2012; Shtulman & Calabi, 2012; To, Tenenbaum, & Hogh, 2016). For example, essentialist ideas reflect implicit beliefs that biological category members all share an underlying and immutable nature that causes their observable properties and makes them what they are. Essentialist reasoning may therefore promote incorrect assumptions that species members do not vary and that species do not change (Samarapungavan & Wiers, 1997; Shtulman & Schulz, 2008; Sinatra, Brem, & Evans, 2008; see also Emmons & Kelemen, 2015). Teleological ideas reflect preferences for purpose-based explanations. They appear to support an early developing tendency to reason that animals' traits originated in order to perform beneficial functions (Atran, 1996; Keil, 1989; Kelemen, 1999). This construal may therefore promote incorrect ideas that adaptations result from purposeful events that uniformly transform individual species members in response to need. Although useful in various everyday reasoning contexts, essentialist and teleological thinking is problematic for accurately representing natural selection as a non-goal-directed mechanism that operates on within-species variation—the property of biological populations that enables differential survival and reproduction to occur. Notably, by 7–8 years of age, these potentially independent biases show signs of coalescing: For instance, children of this age are more likely to deny the possibility of within-species variation when animal traits are described in terms of their functional utility (Emmons & Kelemen, 2015).

One implication of these developmental findings is that education standards that recommend delaying comprehensive instruction on natural selection until middle or high school are placing it too late in the curriculum (see Achieve Inc., 2013; American Association for the Advancement of Science, 2009; National Research Council, 2012). While the motivation for delaying instruction is understandable given, for example, limits on children's background biological knowledge (Metz, 1995; Siegal & Peterson, 1999) and deep-seated assumptions about insurmountable information-processing constraints, it may be doing students a disservice. This is because children have natural explanatory drives (Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1989; Wellman & Gelman, 1992) that can lead them to fill in gaps left open by fragmentary instruction with their own theoretical notions informed by common-sense intuitive reasoning biases. What can result are incomplete and typically incorrect understandings of animal trait origins that when left unchallenged over a number of years can entrench and become increasingly intrusive and difficult to overcome as students attempt to learn and mobilize scientifically accurate explanatory alternatives: For example, consistent with errors seen in undergraduates, untutored older elementary school-age children who are asked about biological adaptation already show marked preferences for essentialist and teleological or goal-directed explanations over selectionist accounts (Berti, Barbetti, & Toneatti, 2017; Berti, Toneatti, & Rosati, 2010; Samarapungavan & Wiers, 1997).

To clarify, it is unlikely that these intuitively based misconceptions ever completely vanish as a result of formal instruction since, consistent with dual processing or conceptual co-existence accounts (e.g., Evans & Stanovich, 2013; Kahneman, 2011; Kelemen & Rosset, 2009; Shtulman & Valcarcel, 2012; see also Legare, Evans, Rosengren, & Harris, 2012), even highly educated adults such as professional scientists default to them when reasoning under duress (e.g., Kelemen, Rottman, & Seston, 2013; Shtulman & Harrington, 2016; see also Järnefelt, Canfield, & Kelemen, 2015). However, even if such commonsense misconceptions are never entirely revised or replaced, the proposal here is that interventions that are initiated from an earlier point may increase the likelihood that students can elaborate scientifically accurate alternatives. This is because they are less likely to filter instruction through ingrained intuitive ideas. Furthermore, having been introduced to correct scientific ideas earlier on, students gain more opportunities to build the representational strength of these ideas which, in turn, increases the likelihood that they will activate these scientifically accurate alternatives during scientific problem-solving. (Emmons, Smith, & Kelemen, 2016; Kelemen, Emmons, Seston-Schillaci, & Ganea, 2014; see also Berti et al., 2017; Kelemen, 2012; Nadelson et al., 2009; Shtulman, Neal, & Lindquist, 2016, for related arguments).

In light of this proposal, recent research has begun to explore the viability of teaching young children a simplified yet comprehensive explanation of natural selection utilizing picture storybooks: This kind of explanation provides a child-friendly account of adaptation by natural selection that is nevertheless complete in that it weaves together several interrelated and complex concepts (e.g., trait variation, inheritance, differential survival, differential reproduction) in a non-fragmentary way to mechanistically explain how a population changes over time. The storybook format was selected for numerous reasons (see Emmons et al., 2016; Kelemen et al., 2014, for detailed explanation). One notable reason is that approaches based in narrative text have previously proven successful in teaching young children various biological facts, including facts and individual concepts pertinent to understanding evolutionary processes (e.g., Brown & Kane, 1988; Brown, Kane, & Long, 1989; Browning & Hohenstein, 2013; Ganea, Ma, & DeLoache, 2011; Legare, Lane, & Evans, 2013; Shtulman et al., 2016; see Dickes & Sengupta, 2013; Horwitz, McIntyre, Lord, O'Dwyer, and Staudt, 2013, for simulation approaches containing some narrative elements). With the goal of using narrative to more comprehensively teach selectionist theory rather than a set of component facts, Kelemen et al. (2014) recently explored whether a picture storybook narrative could lead children to self-generate a detailed, causally coherent mechanistic explanation of adaptation by natural selection absent any overt misconceptions.

Across two studies, Kelemen et al. (2014) presented 5- to 8-year-old children with a custom-made factual narrative picture storybook explaining how a particular trait came to predominate in a phenotypically variable population of realistic but fictitious mammals ("pilosas"). Employing child-friendly language, the image-supported narrative gradually wove together several biological concepts to explain how, over successive generations, the process of differential survival and reproduction led to changes in trait frequencies within the population. Weaving individual concepts together in this way resulted in a cohesive mechanistic explanation that was less prone to misinterpretation because cause-and-effect facts (e.g., food leads to health; health leads to energy; energy leads to survival and fecundity) built on each other page-by-page in a tightly integrated logical causal sequence. The narrative described how the pilosa species went from being a population of animals with predominantly wider trunks to one with predominantly thinner trunks because of selection pressures related to climate change, which prompted their widely available insect food source to descend into thin underground burrows.

Because it was highly uncertain that this initial storybook approach would have any success given conventional educational wisdom about the appropriate timing of natural selection instruction and young children's known information-processing limitations (Bjorklund, 2005; Friedman, 1977), Kelemen et al. (2014) examined children's learning in an optimal setting—a distraction-free research lab where children were read the storybook individually. Using a pretest to post-test design, children were individually evaluated on both their understanding of the isolated facts needed to support an understanding of natural selection (e.g., inheritance, trait constancy) and their learning and generalization of the theory as a coherent whole, including whether they held misconceptions.

Results from the two studies revealed that both younger and older children benefitted from the one-on-one storybook reading (Kelemen et al., 2014). Specifically, although most children could not explain adaptation at pretest, many 5- and 6-year-olds and most 7- and 8-year-olds were able to articulate the selectionist population-based logic of adaptation by natural selection without stating a misconception at post-test. While the first study revealed that children's learning endured following a 3-month delay, the second study demonstrated that children were capable of acquiring an even more nuanced understanding of the multi-step causal logic of natural selection. Across both studies, 7- and 8-year-olds were particularly adept at incorporating the concepts of reproduction and trait frequency changes over multiple generations in their selectionist explanations.

To explore whether these learning gains could be achieved in a more naturalistic setting with a more socio-economically, ethnically, and linguistically diverse population of children, Emmons et al. (2016) then tested children in the more school-like, busy setting of urban after-school programs. Findings from this study revealed that this sample entered the study displaying substantially less background biological knowledge than the lab-based children in Kelemen et al. (2014). Nevertheless, children's learning gains were comparable: Most 5- and 6-year-olds acquired the biological facts needed to support an understanding of natural selection while most 7- and 8-year-olds acquired both the facts and a generalizable selectionist explanation of adaptation by natural selection without stating a misconception.

Based on these studies, early elementary school-age children therefore appear capable of acquiring a comprehensive, coherent understanding of adaptation from an intervention constructed around one causally cohesive, mechanistic storybook. In the present study, we therefore built upon these findings and expanded the intervention: Following Brown et al. (1989) and Gentner, Loewenstein, and Thompson (2003), we employed two analogous storybooks about adaptation and explored young children's capacities to more broadly generalize the mechanism of natural selection across a wider array of transfer scenarios. We also explored whether learning and specifically far generalization would occur when instruction took place in a classroom group rather than in the more optimal one-on-one setting of prior research.

The current focus on far generalization was prompted because while prior studies found that 7- and 8-year-old children displayed pronounced abilities to transfer their learning of natural selection to novel scenarios (Emmons et al., 2016; Kelemen et al., 2014)—contexts that differed substantially from the storybook in terms of surface features (i.e., habit, species type, trait of interest)—the post-test generalization assessment scenarios nevertheless maintained deep structural alignment to the book: Like the book, they focused on traits that somehow related to gaining access to food via foraging (e.g., neck or leg length, beak size). The present investigation therefore moved beyond such structurally aligned scenarios to broader cases, specifically predation scenarios involving adaptations of camouflage-relevant traits (see Shtulman et al., 2016, for related work).

The present intervention was expanded from one to two storybooks because results from the intervention conducted in after-school programs led us to anticipate that far generalization could be a challenge for kindergartners, especially given the greater distractions of a classroom learning context (see Emmons et al., 2016). Furthermore, prior work has demonstrated that the practice of explicitly comparing examples promotes near and far generalization of biological concepts (e.g., Brown & Kane, 1988; Brown et al., 1989; see also Rittle-Johnson & Star, 2011, for evidence in other domains). Thus, in addition to building in a second storybook to underscore the generalizability of the described mechanism and aid abstraction of the selectionist logic of natural selection, we also added a guided classroom discussion that encouraged children to compare the two books. Short- and long-term transfer were examined by including near and far generalization post-tests that were administered immediately following the storybook readings and after a 1-month delay.

Method

Participants

Participants were kindergartners (12 boys, 8 girls, *M* age = 6 years, 2 months, SD = 4 months) and second graders (9 boys, 8 girls, *M* age = 8 years, 5 months, SD = 5 months) from two classrooms within a New England public elementary school. Four additional second graders were tested, but their data were excluded: two children missed the first storybook reading, one failed comprehension check questions during introductions to the assessments, and one participated in an earlier version of the study. Classrooms represented diverse racial, ethnic, socioeconomic, and language backgrounds: 66% of students at the school identified as Hispanic, 22% African American/Black, 8% White, 2% Asian, 2% multi-race or non-Hispanic, and <1% Native American, and 85% of students were eligible for free or reduced price lunch.

Materials and Procedure

Study Environment. Storybook readings and subsequent group discussion of the books took place in the classrooms. Readings and classroom discussions employed a seating arrangement that was customary for classroom read aloud sessions: Children formed a semi-circle on the floor around the experimenter who sat in a chair. After each storybook reading, the classroom resumed its normal activities while children were seen individually for assessments of their understanding of natural selection. Consistent with prior studies, assessments were conducted by members of the research team and took place at a table in the hallway outside of the classrooms behind a small tri-fold privacy screen. As in Emmons et al. (2016), there were moderate levels of background noise due to hallway and nearby classroom activity.

Storybooks. Two custom-made picture storybooks were used in the intervention (see Kelemen & The Child Cognition Lab, 2017, 2018). Custom books were developed because trade books that combine pictures with accurate, comprehensive mechanistic descriptions of adaptation by natural selection were not available at the time of the investigation To eliminate the possibility of differences in baseline knowledge about familiar animals, the two books described realistic but fictitious mammal species. Realistic hand-drawn color illustrations were employed to facilitate children's construal of the animals, habitats, and events described in the books as real. The illustrations were intentionally kept simple and devoid of unnecessary detail or garish color so that the pictures supported, rather than distracted from, the multi-step causal explanation described in the narrative text (DeLoache, 2004; Tare, Chiong, Ganea, & DeLoache, 2010). After all testing was completed, children were debriefed and told that the animals were fictional but that the things that happened in the books happen to real animals and that the process is called natural selection.

Across 12 pages, both books explained biological adaptation by weaving together seven key biological concepts: (i) trait variation inherent to a biological population; (ii) ecological habitat and food-source change due to climate change; (iii) differential health and survival due to differential access to food; (iv) differential reproduction due to differential health; (v) the reliable transmission of heritable physical traits across generations; (vi) the stability and constancy of inherited traits over the lifespan; and (vii) trait-frequency changes (i.e., adaptation) over multiple generations. The gradual nature of adaptation was made salient by the text and the visual display of trait frequency changes within the population across several pages (see Table 1 for more details of the storybook content).

The first book in the sequence was the pilosa storybook used in Experiment 2 of Kelemen et al. (2014; see also Emmons et al., 2016). Consistent with structure mapping theory (Gentner, 1983, 1989; Gentner et al., 2003), the second storybook was designed to strongly align with the relational structure of the first book in order to promote young children's abilities to analogically map, abstract, and generalize the mechanism of adaptation. Thus, it described a novel mammalian species ("dormits") that, like the pilosas, underwent climate change-induced adaptation of a foraging-relevant trait. Despite this similarity, however, the surface structure of the second book differed from the first book in numerous ways: The dormits were highly perceptually distinct from the anteater-like pilosas by virtue of being bipedal mammals with physical similarities to a mongoose or kangaroo rat. Their diet and habitat also differed. Rather than eating insects, they ate nuts that grew mostly on the uppermost branches of trees after climate change rendered their initially sunny wooded environment permanently icy and cold. In this new habitat, dormits with longer backs had greater fitness than those with shorter backs because they could reach the higher branches that now contained most of the nuts. An additional page was included at the beginning of the book to highlight that the depicted adult dormits varied in height by virtue of inherent variability within the adult population and not because of differences in age or developmental growth. This step was taken because piloting indicated that, consistent with an essentialist view, children tended to dismiss the height variations as mere reflections of maturity differences rather than as true adult phenotypic variability (see Herrmann, French, DeHart, & Rosengren, 2013; Rosengren, Gelman, Kalish, & McCormick, 1991, for children's intuitive theories about growth).

The two books were read 1 week apart and each read-through was entirely monological to ensure undistracted focus on the multi-step causal explanation as it was laid out page-by-page. Prior to the second storybook reading, the experimenter briefly recapped what happened in the first storybook and gave children the following prompt to align and analogically compare the examples in order to aid them in abstracting the common underlying structure: "Today, we're going to read *The Story of the Dormits*. The dormits look really different than the pilosas. But actually, what happens to the pilosas and what happens to the dormits is almost the same. While I read you *The Story of the Dormits*, I want you to listen carefully and think about what makes living and surviving hard for some of the dormits and their children."

Analogical Classroom Discussion. After hearing the second storybook, children engaged in a structured experimenter-led classroom discussion explicitly focused on drawing analogies between the two books. The goal of this activity was to encourage children to process the deep-level relational similarities of the two books. Specifically, it aimed not only to promote children's ability to engage in near transfer (i.e., by applying their learning to other distinct foraging contexts) but also their ability to abstract the underlying mechanism for purposes of far transfer (i.e., by extending their learning to scenarios involving camouflage in predator-prey contexts; see Brown & Kane, 1988; Brown et al., 1989).

During the discussion, the experimenter posed questions that prompted children to think about similarities between the two book narratives with respect to: (i) within-species variation in the past and present; (ii) climate change; (iii) differential access to food; (iv) differential survival; (v) differential reproduction; and (vi) multiple generations. When a child gave a correct response to a question, the experimenter confirmed the answer by saying, "That's right." If a child offered an incorrect answer, which occurred rarely, the experimenter naturalistically gave other children opportunities to respond and offer a correct answer (e.g., by saying, "Does anyone else have an idea about that?"). This approach allowed us to remain largely consistent with our pre- and posttest

Table 1
Summary of pilosa storybook images and narrative content

Page number	Pictures	Narrative content[a]
1 (Title Page)	Past population of pilosas with children in past habitat.	Title: *The Story of the Pilosas.*
2	Past and contemporary populations of adult pilosas on white background.	Description of how the population of pilosas looked in the past and how it looks now. Description emphasizes trait variation and the frequency of animals with each trait variation (wider vs. thinner trunks) in the past and contemporary populations. Asks, "Why do pilosas mostly have thinner trunks now?"
3	Past population of adult pilosas with children in past habitat (food source shown); wider-trunked pilosas depicted with more children.	Description of how the pilosas lived many hundreds of years ago, including where they lived and how they spent time looking for food and water.
4	Past population of adult pilosas with children in past habitat (food source shown); wider-trunked pilosas depicted with more children.	Description of what the pilosas ate (milli bugs) and where the bugs lived. Cause-and-effect description of how when the pilosas ate the bugs they got healthy and strong, and healthy pilosas had energy to have many children.
5	Staggered images of (i) past habitat and (ii) contemporary habitat (food source shown).	Description of how the weather changed and became very hot, and as a result, most of the bugs moved to underground tunnels.
6	Past population of adult pilosas in contemporary habitat (food source shown)	Description of how the small number of pilosas with thinner trunks could eat lots of bugs because their trunks could fit into the underground tunnels, whereas, pilosas with wider trunks struggled to find food. Description of how some pilosas with wider trunks could eat, but some did not.
7	Past population of adult pilosas in contemporary habitat (food source shown).	Description of how the small number of pilosas with thinner trunks were strong and healthy because they were able to eat lots of bugs. Description of how some pilosas with wider trunks were kind of healthy but others were not because they couldn't get any bugs. These pilosas were weak and didn't have any energy.
8	Past population of adult pilosas with children in contemporary habitat (food source shown); thinner-trunked pilosas depicted with more children.	Description of how the small number of pilosas with thinner trunks were very healthy and so lived for a long time and had enough energy to have many children. Description of how their children were born with thinner trunks because children usually resemble their parents. Description of how some pilosas with wider trunks were so weak that they died before having children, and how other pilosas with wider trunks were only healthy enough to have one child that was born with a wider trunk like its parents.

continued

Page number	Pictures	Narrative content[a]
9	Staggered images of (i) original past population of adult pilosas with children in contemporary habitat and (ii) the next generation of adult pilosas (the grown children) in contemporary habitat.	Description of how many years went by and the pilosas who had lived long enough to have children got old and died. Description of how their children grew up and still had the same trunk they were born with because that stayed the same their whole life. Description of how the population of pilosas now looked different. There were still a lot of piloas with wider trunks but because pilosas with thinner trunks had been healthier and had many children, there were more pilosas with thinner trunks than before.
10	Staggered images of (i) the next generation of adult pilosas and their children in contemporary habitat and (ii) the third generation of adult pilosas (the grown children) in the contemporary habitat.	Description of how time passed and the same events happened again. Summary of how pilosas with thinner trunks got more food, were healthier, and had more children born with thinner trunks. Although there were still many pilosas with wider trunks, they were less healthy and had fewer children. As a result, the population looked different such that there were more pilosas with thinner trunks than with wider trunks.
11	Staggered images of (i) the third generation of adult pilosas and their children in the contemporary habitat and (ii) the fourth generation of adult pilosas with their children in the new habitat.	Description of how this process happened again and again. Summary of how pilosas with thinner trunks got more food, were healthier, and had more children born with thinner trunks. These children grew up and had lots more children. Pilosas with wider trunks got less food, were less healthy, and had fewer children. Over many cycles, there came to be more pilosas with thinner trunks in the population.
12	Past and contemporary populations of adult pilosas on white background.	Summary statement concluding that how pilosas went from being a population of animals with mostly wider trunks to one with mostly thinner trunks has now been explained.

[a] Narrative content has been simplified for the purposes of this table and does not constitute verbatim content. For the full storybook, see Kelemen and The Child Cognition Lab (2017).

assessment approach in which children were not given explicit, direct corrective feedback. To ensure that the amount and content of instruction was the same across both classrooms, for each prompt, the experimenter gave a scripted response echoing a correct answer already provided by a student (See Supplementary Materials for full analogical prompts).

The classroom discussion closed with the experimenter asking children to consider how the books' events might apply to other animals with different types of body parts. This was done to encourage children to consider the generalizability of the mechanism in the books. In this part of the discussion, children were asked to reason about two novel scenarios of adaptation. Namely, they were asked to judge outcomes for species members who possess either the advantageous or disadvantageous trait in an environment where water access was difficult or where there were poisonous bugs. After the class responded, the experimenter again echoed back the correct response. While these examples were intended to scaffold reasoning about adaptation beyond the foraging scenarios narrated in the storybooks, importantly, neither discussion example involved a camouflage-relevant trait or a predation context. These were reserved for the far generalization trials.

Assessment. Children's understanding of natural selection was evaluated a total of seven times across four different testing days using parallel assessments that combined open-ended questions with closed-ended questions and requests for justifications. Children's opportunities to generate justifications and explanations during these assessments were viewed as an intrinsic part of their learning, consistent with research on testing and self-explanation effects (e.g., Chi, De Leeuw, Chiu, & LaVancher, 1994; Rittle-Johnson, 2006; Roediger & Karpicke, 2006; Williams, Lombrozo, Rehder, 2013). Importantly, however, while the repeated opportunity to explain had the potential to benefit children by reinforcing and extending correct knowledge, it could also hinder them. This is because corrective feedback was never offered at any point in any assessment. Therefore, children who articulated a misconception ran the risk of repeatedly explaining the same mistaken logic across all assessments.

Intervention Timetable. Figure 1 summarizes when each portion of the intervention occurred over a 7-week period. Briefly, the pretest assessment evaluated children's knowledge of natural selection before the classroom intervention. After the first storybook reading occurred, children's understanding of natural selection was evaluated twice: A comprehension post-test evaluated their understanding of the pilosa book (book 1 comprehension) and a generalization post-test examined their ability to apply their learning to another novel species that underwent adaptation on a foraging-relevant trait (book 1 near generalization). After the second storybook and classroom discussion, children were again assessed twice, first within the context of another novel foraging-relevant trait (book 2 near generalization) and then within the dissimilar context of a camouflage-relevant trait (book 2 far generalization). The final day of testing examined enduring learning 1 month later: Children were again assessed on their understanding of natural selection in a novel

Week 1	PRETEST: foraging trait, novel animal
Week 2	BOOK 1 READING: *The Story of the Pilosas* BOOK 1 COMPREHENSION: foraging trait, book 1 animal (pilosas) BOOK 1 NEAR GENERALIZATION: foraging trait, novel animal
Week 3	BOOK 2 READING: *The Story of the Dormits* ANALOGICAL CLASSROOM DISCUSSION BOOK 2 NEAR GENERALIZATION: foraging trait, novel animal BOOK 2 FAR GENERALIZATION: camouflage trait, novel animal
Week 4 *Week 5* *Week 6*	1-MONTH DELAY: No interaction with experimenters
Week 7	DELAYED NEAR GENERALIZATION: foraging trait, novel animal DELAYED FAR GENERALIZATION: camouflage trait, novel animal

Figure 1. Two-storybook classroom intervention timetable.

foraging context (delayed near generalization) and in a novel camouflage context (delayed far generalization).

Another feature of the timetable was the deliberate placement of assessment animals designed to be increasingly difficult as the intervention progressed. First, the pretest and book 1 near generalization assessments (counterbalanced) utilized species whose focal trait did not create a height disparity between species members (i.e., arm, tail length): This echoed the kind of trait introduced in the pilosa storybook (trunk width). These conceptually easier cases were introduced early in the intervention to encourage acceptance of within-species variation and counteract the essentialist view that members vary in appearance solely because they are at different stages of development. Second, to evaluate learning in a more difficult context where growth-based misinterpretations of variation were more conceptually available, book 2 and delayed near generalization assessments (counterbalanced) were introduced next. These assessments employed animals that, like the dormits, varied on a focal trait yielding a height disparity between members (neck, leg length): These cases could be understood in terms of a developmental misconception, rooted in essentialist notions, that species acquire beneficial traits through growth. Other misconceptions that were possible included explicit transformationist views of species change (e.g., ideas that animals transform in response to need) and “bigger is better” notions (e.g., ideas that larger animals have a natural advantage; Silvera, Josephs, & Giesler, 2002). Finally, the book 2 and delayed far generalization assessments (counterbalanced) introduced a conceptual challenge beyond the novel camouflage context: In these assessments, it was animals with the diminished form of the trait, rather than the amplified form, that were better able to hide from their natural predator following climate change. Prior work with adults has found that these kinds of relaxed adaptation cases where traits reduce or disappear are particularly challenging potentially because they run counter to teleological misconceptions of evolution as a goal-directed perfecting force (Ha & Nehm, 2014).

Assessment Questions. The questions used in each assessment were identical to those used in prior work (Emmons et al., 2016; Kelemen et al., 2014, Experiment 2).[1] For each assessment species, children were talked through two pairs of images to provide the setup for subsequent questions. One pair depicted the past appearance of the species in which a particular variant of a focal trait predominated (e.g., longer ear length) and the species’ past habitat including either its food source or natural predator (e.g., aerial predator within a tall grass habitat where the animals hid). The other pair depicted the species’ present habitat post-climate change including either its current food source or natural predator (e.g., aerial predator within a shorter grass habitat where the animals hid) and the present appearance of the species in which a different variant of the focal trait predominated (e.g., shorter ear length). Children had to infer that the trait of interest was pertinent to either gaining access to food or being detected by a predator because information about trait functions or affordances was never explicitly stated (see Emmons & Kelemen, 2015, for further discussion).

Following each setup, children answered six closed-ended questions that evaluated their knowledge of a series of isolated facts that support a correct understanding of natural selection (see Table 2). Questions were asked while children viewed relevant images of the species and tapped four adaptation concepts: (i) differential survival (2 questions); (ii) differential reproduction (2 questions); (iii) constancy of traits over the lifespan (1 question), (iv) inherited family resemblance (1 question). Children were required to justify their answers, and it was this

Table 2

Closed-ended isolated fact questions with sample justifications

Concept	Trait type	Question	Accurate justification	Inaccurate justification
Differential Survival (advantaged member)	Foraging	Nowadays, will a *tardon with a stretchier tail* probably be healthy and live for a long time? Why/Why not?	Yes, they could eat more melons and get a lot of energy and stay alive.	Yes, because they are younger.
Differential Survival (disadvantaged member)	Camouflage	Nowadays, will a *tomad with a longer tail* probably be healthy and live for a long time? Why/Why not?	No, because their longer tail can stick out and these [points to predator] can eat them.	Yes, because it can get a lot of food.
Differential Reproduction (advantaged member)	Foraging	Nowadays, will a *tardon with a stretchier tail* probably have lots of children? Why/Why not?	Yes, because they eat lots of food and they will be healthy.	Yes, because they live nowadays.
Differential Reproduction (disadvantaged member)	Camouflage	Nowadays, will a *manu with longer ears* probably have lots of children? Why/Why not?	No, because the [points to predator] will eat him and they won't be able to hide.	Yes, because it is an adult.
Trait Knowledge: Inheritance	Foraging	These fully-grown *wilkies* both have *shorter legs*. If these two *wilkie*s had a child, what kind of *legs* [*longer or shorter*] would their child probably have? Why?	Shorter, because it is gonna have the same legs as its parents.	Shorter, because when they were babies they had shorter legs.
Trait knowledge: Trait constancy	Camouflage	See this young *manu*. It was born with *longer ears*. When this *manu* is fully-grown, will it be an adult with *longer ears or an adult with shorter ears?* Why?	Longer, because how it's born with, it's going to stay like it.	Longer, because this one [points to shorter-eared adult] is too short.

Note: Italicized information differed depending on the animal species under consideration. For the differential survival and reproduction questions, the presentation of advantaged and disadvantaged animals was counterbalanced.

justification—not the accuracy of closed-ended answers—that determined whether they were given credit for understanding the concept.

Following the isolated fact questions, children were asked three open-ended questions (e.g., "How do you think that (change in trait frequency) happened?") to probe whether they could self-generate a correct explanation of population change in terms of natural selection (see Table 3).

Table 3
Open-ended questions

Questions
Many hundreds of years ago, most of the *fully-grown pilosas had wider trunks*, but now most of the *fully-grown pilosas have thinner trunks*. How do you think that happened?
What happened to *pilosas with thinner trunks*? Why?
What happened next after…? [E repeats P's response to previous question] Why?
What happened next after…? [E repeats P's response to previous question] Why?
What happened to *pilosas with wider trunks*? Why?
What happened next after…? [E repeats P's response to previous question] Why?
What happened next after…? [E repeats P's response to previous question] Why?

Note: Italicized information differed depending on the animal species under consideration. Question orders about advantaged and disadvantaged animals were counterbalanced. E, experimenter; P, participant.

While responding to these questions, children viewed images of the past and present populations. In order to determine the extent of their understanding, children received a standard set of prompts designed to encourage them to elaborate on their reasoning. These prompts took the form of the experimenter repeating back what the child had already said (e.g., "What happened next after [child's previous response]?") and were necessary given children's tendencies to truncate their answers. By virtue of the experimenter never providing new information or correcting children, these requests had the potential to reveal misconceptions masked by children's unelaborated responses or, alternatively, to uncover their more sophisticated understanding of natural selection.

Coding. To facilitate comparisons with prior work, the coding scheme from earlier studies was used (Emmons et al., 2016; Kelemen et al., 2014, Experiment 2). For each test, children were assigned a global score that captured their natural selection understanding across responses to all closed- and open-ended questions. Responses were coded based on a conceptual checklist and conservative coding rubric (see Table 4 for details and sample responses). Using a global measure permitted examining whether children's overall theoretical understanding changed over the course of the intervention. This coding approach differed from approaches in prior work that have focused on children's learning of the constituent elements of natural selection understanding rather than the connection between concepts and their learning of the theory as a whole (e.g., Browning & Hohenstein, 2013; Legare et al., 2013; Shtulman et al., 2016). We utilized a global score because while it is valuable to assess whether children, on average, can learn individual evolutionary sub-concepts, a more granular piece-by-piece approach runs the risk of crediting children who adopt a blended account or display a misconception with an accurate causal understanding of natural selection. Because the goal of the current research was to assess whether young children can elaborate and generalize an accurate and coherent selectionist account—contrary to assumptions embodied in current educational guidelines——we therefore chose to avoid this risk. Coding criteria therefore focused on evaluating the causal cohesiveness, completeness, and accuracy of each child's understanding of adaptation at each assessment. Children displaying any overt indication of a misconception were thus never credited with mobilizing a correct theoretical understanding of natural selection.

Children's understanding of natural selection was classified into one of six hierarchical levels. Their understanding was categorized as Level 0, "no isolated facts," when responses to the closed-ended questions demonstrated limited or no knowledge of the prerequisite facts needed to support an understanding of natural selection. Consistent with a conservative evaluation approach,

Table 4
Conceptual checklist of NS understanding and sample partial open-ended responses

Levels and checklist	Partial open-ended responses following one or more of the three open-ended questions[a]
Level 0: No isolated facts <3 closed-ended questions correctly answered and justified.	Not applicable
Level 1a: Partial facts but no NS understanding Between 3 and 4 closed-ended questions correctly answered and justified.	Not applicable
Level 1b: Isolated facts but no NS understanding Meets criteria for isolated facts (5 out of 6 closed-ended questions correctly answered and justified) but one or more of the following is present: (i) Misconception; (ii) No mention of differential survival advantage; (iii) Inaccurate mention of differential survival or reproduction.	*Example of a misconception*: E: Many hundreds of years ago most of the fully-grown orpeds had shorter arms but now most of the fully-grown orpeds have longer arms. P: Cause they all grew up. E: How do you think that happened? P: Cause they all grew up, like the mom or the dad. *Example of inaccurate mention of differential survival*: E: What happened to manus with longer ears? P: They stayed alive. E: Why? P: Because um...because they're healthier and they will have more babies.
Level 2: Foundation for NS understanding All of the following are present: (i) Meets criteria for isolated facts; (ii) No misconception; (iii) Accurate mention of differential survival.	*Example of correct mention of differential survival, but no mention of differential reproduction*: E: What happened to the pilosas with thinner trunks? P: These ones? E: Yeah. What happened to them? P: Um, they stayed alive. E: Why? P: Cause got longer noses. E: Okay. P: They, they just put their noses in there and eat it. E: Oh, okay. And so what happened next after they stayed alive? P: Um...they died.
Level 3: NS understanding in one generation All of the following are present: (i) Meets criteria for isolated facts; (ii) No misconception; (iii) Accurate mention of differential survival; (iv) Accurate mention of differential reproduction in one generation.	*Example of correct mention of differential survival and differential reproduction in one generation:* E: What happened to the rudoos with longer necks? P: They get to have babies. E: Why? P: Because they ate...um...they ate a lot and they got healthy and strong. E: Oh, and what happened next after they got to have babies? P: The babies had the same neck as them [points to long-necked rudoos] and then they got to eat...um...a lot and then they died.
Level 4: NS understanding for multiple generations All of the following are present: (i) Meets criteria for isolated facts; (ii) No misconception; (iii) Accurate mention of differential survival; (iv) Accurate mention of differential reproduction	*Example of correct mention of differential survival and differential reproduction in multiple generations*: E: Oh, so what happened to the wilkies with longer legs?

continued

Levels and checklist	Partial open-ended responses following one or more of the three open-ended questions[a]
in one generation; (v) Accurate mention of differential reproduction in multiple generations.	P: They can eat a lot of berries. E: Oh, why? P: They have long legs, they can eat a lot of berries. E: Oh, and what happened next after they could eat a lot of berries? P: When they grow up, they will have a lot of childs. E: Oh, why? P: Because they will, because they...um. . .they could eat a lot because they have long legs. E: Oh, and what happened next after they had a lot of children? P: They died and then the children, they get to eat a lot and then and then they have more children, then more and more. E: Oh. I see. And so what happened to the wilkies with shorter legs? P: They can only have at most one child because they can. . .um...only eat a little bit of lemons. E: Oh, and why's that? P: Because they have short legs, they can't eat a lot of lemons. E: Oh. What happened next after they couldn't eat a lot? P: They died and they can only have. . .um. . .one children. E: Oh, why? P: Because they couldn't eat a lot.

Note: NS, natural selection; E, experimenter; P, participant. Italicized information differed depending species.

[a] The initial open-ended question, "How do you think that (population change) happened?" was followed by subsequent requests for elaboration (see Table 3). Sample responses reported here have been edited for length such that only partial responses, often in connection to requests for elaboration, are shown to illustrate specific concepts that were coded as part of the conceptual checklist. These sample partial responses do not reflect any one child's entire open-ended response.

children who did not provide correct responses to at least three of the closed-ended questions fell into this category even if they offered accurate responses to open-ended questions. Understanding was categorized as Level 1a, "partial facts but no natural-selection understanding," when children correctly answered between three and four of the closed-ended questions. This Level 1 subcategory (Level 1a) was not present in Kelemen et al. (2014) or Emmons et al. (2016); however, it was added as a separate analytic level in the current research to capture learning at a finer grain given the challenges of classroom group learning over individual learning. Level 1b, "isolated facts but no natural-selection understanding," was reserved for children who displayed robust factual knowledge: Children who correctly answered between five and six of the closed-ended questions yet did not demonstrate a correct understanding of adaptation by natural selection in their responses to the open-ended questions were assigned to this category.

The three highest levels of understanding (Levels 2–4) were only assigned when children demonstrated a robust understanding of the isolated facts (i.e., correctly answering ≥ 5 closed-ended questions) and provided a correct self-generated explanation of the selectionist logic of adaptation in response to open-ended questions. Children were only ever assigned a Level 2 or higher categorization if there was no sign of a misconception at any point in the assessment.

Children were assigned Level 2, "foundation for natural-selection understanding," when their open-ended responses focused on adaptation resulting from differential survival, namely the idea that species members with disadvantageous traits often died while those with advantageous traits tended to survive as a result of selection pressures. They were assigned Level 3, "natural-selection understanding in one generation," when they explained adaptation both in terms of differential survival and differential reproduction but their open-ended explanations were limited to considering these processes only in the initial population following climate change and their immediate descendants. Finally, children were assigned Level 4, "natural-selection understanding for multiple generations," when their open-ended explanations were expanded to explicitly reference the concept that adaptation occurs over multiple generations. Interrater reliability between two coders was excellent ($\kappa = 0.89$), and all disagreements were resolved through discussion.

Results

Following earlier studies (Emmons et al., 2016; Kelemen et al., 2014), data were analyzed using repeated measures ordinal logistic regressions. This analysis examined how the distribution of children across the six hierarchical levels of natural selection understanding changed across the seven assessment times. Odds ratio statistics further indicated the magnitude of change in the odds that children's understanding of natural selection improved by one or more levels between two specific assessment times.

Kindergartners

Repeated measures ordinal logistic regressions, factoring in all seven assessment times, revealed that the two-storybook classroom intervention induced learning among kindergartners, Wald $\chi^2(6) = 57.82, p < 0.001$ (Figure 2).

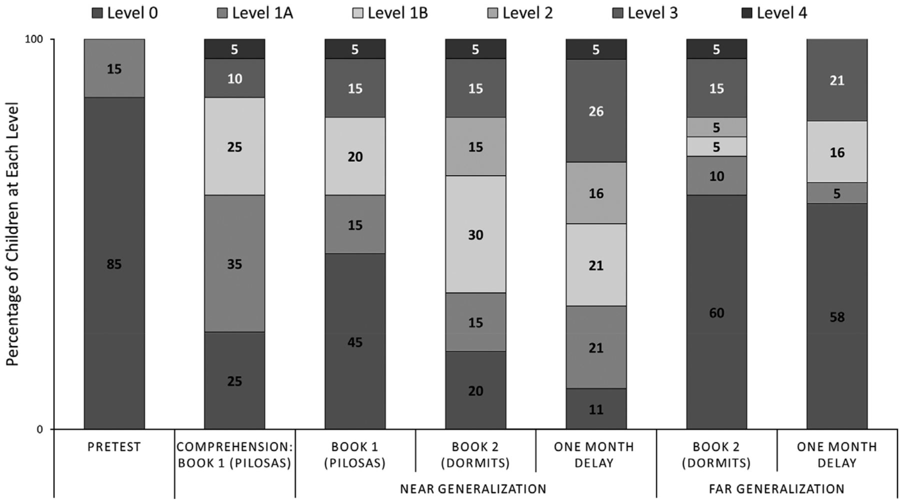

Figure 2. Percentages of kindergartners classified into the six levels of natural selection understanding as a function of assessment. Because of rounding, percentages do not always add up to 100%. Level 0, no isolated facts; Level 1a, partial facts but no natural-selection understanding; Level 1b, islolated facts but no natual-selection understanding; Level 2, foundation for natural-selection understanding; Level 3, natural-selection understanding in one generation; Level 4, natural-selection understanding for multiple generations. [Color figure can be viewed at wileyonlinelibrary.com].

Learning Following the First Storybook. Given kindergartners' starting levels of understanding at pretest, their odds of being in a higher level of natural selection understanding following the first storybook reading increased eleven-fold, OR = 11.68, $p = 0.001$, 95% CI (2.88, 47.32). This improvement was primarily due to learning the facts prerequisite to understanding natural selection. Specifically, while 85% of kindergartners were in Level 0 at pretest because they demonstrated no factual knowledge, after the pilosa story, this pattern changed: At book 1 comprehension, 35% of kindergartners demonstrated partial factual knowledge (Level 1a), 25% had robust factual knowledge (Level 1b), and another 15% displayed a correct selectionist understanding of adaptation (Level 2 or higher). Notably, kindergartners were able to successfully generalize what they had learned to a different novel species that underwent adaption on another foraging trait: There was no change in their odds of being in a higher level of natural selection understanding from book 1 comprehension to book 1 near generalization, $p = 0.30$. This indicated that their learning was stable and transferrable.

Near Generalization Following the Second Storybook and Classroom Discussion. Kindergartners also benefitted from the second storybook and analogical classroom discussion. Using book 1 near generalization as the baseline, their odds of being in a higher level of understanding at book 2 near generalization increased two-fold, OR = 2.37, $p = 0.02$, 95% CI (1.18, 4.77). These gains reflected further increases in their factual knowledge as well as their understanding of natural selection as a theoretical whole: The percentage of children with no factual knowledge (Level 0) dropped from 45% to 20%, whereas, the percentage who could give an accurate (at minimum Level 2) selectionist explanation of natural selection increased from 20% to 35%. Importantly, kindergartners' learning endured following a one-month delay: At delayed near generalization, 47% of children achieved the three highest levels of natural-selection understanding. Although there was a trend suggesting that performance actually increased between book 2 generalization and delayed near generalization, it did not reach significance, $p = 0.09$.

Far Generalization Following the Second Storybook and Classroom Discussion. Despite increases in performance that maintained across comprehension and near generalization tests, consistent with the known challenges of transfer (Brown & Kane, 1988; Brown et al., 1989; Gentner, 1989), kindergartners' performance dipped when they were asked about a non-foraging-based scenario involving predators and adaptation with respect to camouflage traits. They showed a nearly four-fold decrease in their odds of being in a higher level of natural selection understanding between book 2 near and far generalization assessments, OR = 0.26, $p = 0.001$, 95% CI (0.12, 0.58). This performance decline appeared to occur primarily because many kindergartners were unable to correctly apply their factual knowledge in the camouflage context: Only 20% of children displayed no factual knowledge (Level 0) at book 2 near generalization, but this increased to 60% at book 2 far generalization. Nevertheless, kindergartners showed no further performance decline on far transfer following a one-month delay: There was no change in their odds between book 2 far generalization and delayed far generalization, $p = 0.81$. These findings suggest that while kindergartners' learning was stable and transferable in cases of near generalization, they encountered greater difficulty with far generalization.

Second Graders

Repeated measures ordinal logistic regressions showed that the two-storybook classroom intervention also induced learning among second graders, Wald $\chi^2(6) = 120.80$, $p < 0.001$ (see Figure 3).

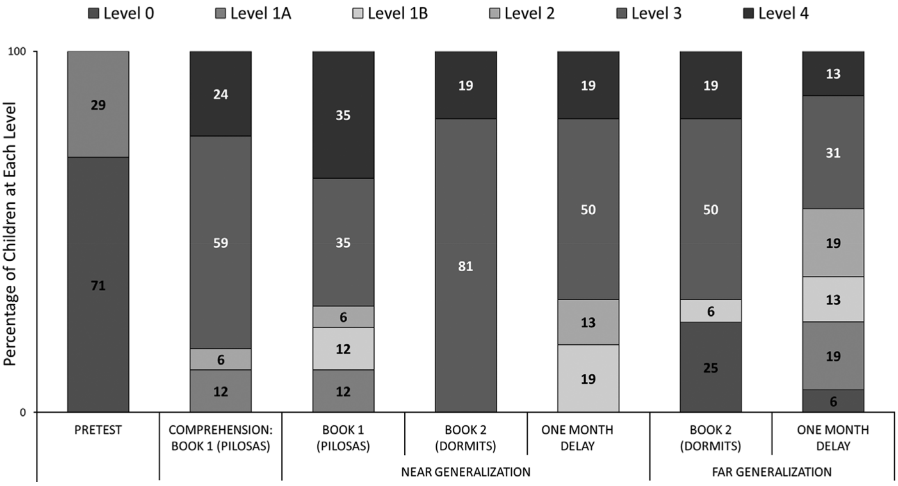

Figure 3. Percentages of second graders classified into the six levels of natural selection understanding as a function of assessment. Because of rounding, percentages do not always add up to 100%. Level 0, no isolated facts; Level 1a, partial facts but no natural-selection understanding; Level 1b, islolated facts but no natual-selection understanding; Level 2, foundation for natural-selection understanding; Level 3, natural-selection understanding in one generation; Level 4, natural-selection understanding for multiple generations. [Color figure can be viewed at wileyonlinelibrary.com].

Learning Following the First Storybook. Second graders' learning following the first book showed marked gains: From pretest to book 1 comprehension, their odds of achieving a higher level of natural selection understanding increased a substantial one hundred twenty-fold, $\mathrm{OR} = 120.74, p < 0.001$, 95% CI (30.03, 485.47). This degree of learning was the result of second graders not only acquiring factual knowledge but also considerable theoretical knowledge about natural selection. Although 71% of second graders were at Level 0 by demonstrating no factual understanding at pretest, after hearing the first storybook, 88% of second graders accurately described natural selection within the pilosas (Level 2 or higher). Indeed, 59% of second graders incorporated the concept of differential reproduction in their responses (Level 3), while another 24% reached the highest level of understanding by explicitly describing adaptation occurring over multiple generations (Level 4). Like kindergartners, second graders were able to generalize what they learned to a new species undergoing adaptation on a foraging trait: They displayed no change in their odds of being in a higher level of natural selection understanding between book 1 comprehension and book 1 near generalization, $p = 0.94$. Therefore, their learning maintained within a near transfer context.

Near Generalization Following the Second Storybook and Classroom Discussion. After the second storybook and classroom discussion, second graders' learning remained robust: At book 2 near generalization, 100% of children performed in the two highest levels of natural selection understanding (Levels 3 and 4). However, given that children's near generalization abilities following the first storybook were already strong—88% achieved Level 2 or higher—this pattern of near ceiling performance meant that there was no significant difference in performance between book 1 and book 2 near generalization assessments, $p = 0.65$. Second graders' learning also endured following a one-month delay: They showed no performance decline between book 2 near generalization and delayed near generalization, $p = 0.26$.

Far Generalization Following the Second Storybook and Classroom Discussion. In contrast to kindergartners, second graders showed robust abilities to abstract the logic of natural selection and engage in far generalization when a species adaptation involved a trait pertinent to camouflage: 69% of second graders achieved the two highest levels of natural selection understanding (Levels 3 and 4) on book 2 far generalization. Indeed, their performance on far generalization did not differ significantly from that on near generalization immediately following the second storybook reading and classroom discussion, $p = 0.12$. Furthermore, their success on far generalization endured a month later. Sixty-three percent of second graders correctly applied their natural selection understanding (Level 2 or higher) to another scenario involving a camouflage trait: There was no change in their odds between book 2 far generalization and delayed far generalization, $p = 0.32$.

Coding of Misconceptions

Consistent with the conservative coding scheme aimed at evaluating the accuracy and cohesiveness of children's understanding of adaptation, individuals who displayed a misconception in response to any question in the assessment were potentially downgraded in their overall level of understanding. Thus, even if children demonstrated many accurate ideas, those who stated a misconception at any point were never credited with having a correct selectionist theory of adaptation (Level 2 or higher). To examine whether the intervention was successful at providing an alternative framework for understanding why species have the specific traits that they do and therefore successfully combat any preexisting misconceptions, the frequency and type of children's misconceptions at each assessment was explored. Misconceptions were coded into several independent categories, but they were not mutually exclusive given that a child could be assigned multiple misconceptions within a single assessment.

Children were assigned a *developmental* misconception if they ever explained the trait variation observed in terms of age differences between the animals. Importantly, during the initial setup, children were told that the species members with trait variations were fully grown and this fact was reiterated at numerous points throughout the assessment. A *transformationist* misconception was assigned if children explained species change in terms of individual species members acquiring an advantageous trait within their lifetime without referencing a specific mechanism. By contrast, children who stated a transformationst idea but also explicitly referred to the animal's need, effort, or a purpose as driving the transformation were assigned a *teleological/need/effort* misconception. When children explained that species with an amplified form of the trait were better able to survive and/or reproduce but did not elaborate further on why that was, they were assigned a *bigger is better* misconception. This type of reasoning displays mistaken causal logic about why an animal is advantaged or disadvantaged and may represent a mental shortcut guided by teleological ideas that nature moves toward perfection or always prefers members with amplified forms of survival-relevant traits (see Ha & Nehm, 2014; Silvera et al., 2002). Any other incorrect idea about trait variation or species change was assigned an *other* misconception. Coders agreed on 95% of cases, and all disagreements were resolved through discussion.

As shown in Table 5, the vast majority of children from both grades came into the study with at least one misconception. The most prominently stated misconception was developmental followed by transformationist. Children's robust tendencies at pretest to interpret trait variation in terms of age-related differences in growth was striking given that the pretest assessment species was purposefully designed to avoid this type of characterization by virtue of the variation in their focal traits (arm, tail length) not yielding an overall height disparity among members. Children also were explicitly told that—despite their variability—all species members depicted were fully grown adults.

Table 5

Percentage of kindergartners and second graders stating misconception(s) about natural selection at each assessment

	Pretest	Book 1 comp.	Book 1 near gen.	Book 2 near gen.	Delayed near gen.	Book 2 far gen.	Delayed far gen.
Kindergartners							
Any misconception	85	50	65	60	42	55	42
Developmental	75	35	55	40	32	45	32
Transformationist	25	15	20	35	11	15	16
Bigger is better	5	5	5	0	0	0	5
Teleo/need/effort	5	0	5	0	5	5	0
Other	5	0	0	0	11	0	0
Second graders							
Any misconception	88	6	24	0	13	6	6
Developmental	71	0	12	0	0	6	0
Transformationist	41	6	18	0	13	0	6
Bigger is better	18	0	6	0	0	0	0
Teleo/need/effort	12	0	0	0	0	0	0
Other	6	0	0	0	0	0	0

Related-samples McNemar tests revealed that the proportion of kindergartners and second graders coded with any misconception was lower at comprehension than at pretest, $p = 0.02$ and $p < 0.001$, respectively: Developmental misconceptions significantly declined for both ages, and transformationist misconceptions declined among older children (sig. ps < 0.05; tests were not run on other misconceptions because they occurred infrequently). Related-samples McNemar tests comparing post-test assessments to one another found no other differences in the proportion of children holding any misconception (ps > 0.05). Based on kindergartners' pattern of showing a higher proportion of misconceptions than second graders on assessments following the second part of the intervention, it appeared that kindergartners' misconceptions were resistant to further revision even after additional instruction. This resistance may have been compounded by the deliberate placement of a more challenging assessment species after the second storybook. Second graders, conversely, rarely defaulted to their earlier misconceptions in post-test assessments even when tested on the most difficult far generalization trials.

Discussion

Findings from the current investigation reveal that early elementary school-aged children are capable of self-generating and generalizing an accurate, cohesive explanation of adaptation by natural selection across a range of diverse scenarios when learning is initiated in a classroom group setting. Children were able to apply their learning across increasingly difficult assessments and following a 1-month delay. Echoing earlier results (Emmons et al., 2016; Kelemen et al., 2014), second graders were particularly successful at generating the selectionist logic of natural selection absent misconception and applying it to varied adaptation contexts. However, while second graders showed the greatest learning, kindergartners also benefitted from the type of intervention used here.

Kindergartners' learning most consistently reflected their acquisition of the isolated facts that support understanding the theory of natural selection. Data on their misconceptions suggested that this emphasis on facts rather the theory as a coherent whole may, in part, have derived from pronounced essentialist beliefs that species are homogeneous—beliefs that were resistant to instruction amid the distractions of the classroom learning setting. Notably, even on the comprehension test for the first storybook—and in contrast to prior one-on-one learning studies (e.g., Emmons et al., 2016)—many kindergartners continued to view trait differences in terms of age-related growth, while others incorrectly interpreted species change as occurring at the individual level. Despite these challenges, a month after hearing the second storybook, nearly half of kindergartners generalized their correct theoretical understanding of natural selection (Level 2 or higher) in a near generalization context suggesting that, given time and opportunity, learning difficulties posed by the essentialist bias may be surmountable by many young children when instruction is provided in a classroom group setting.

Such findings provide evidence for the contention that even very young children are often capable of far more abstract explanatory sophistication than is generally assumed (e.g., Brown et al., 1989; Emmons et al., 2016; Kelemen et al., 2014; Metz, 2009). Nevertheless, it is important to qualify that even with these gains, the scope of kindergartners' generalizable learning was restricted. Specifically, they primarily displayed their generalization abilities on post-test assessments that, like the storybooks, focused on foraging-relevant traits. Several factors might have contributed to this limitation. One possibility is that the focus on foraging in both storybooks and in the analogical discussion derailed young children's far transfer abilities, leading them to inaccurately infer the general causal relevance of food access and to ignore other survival-relevant dynamics (e.g., camouflage-predation). Consistent with this notion, closer scrutiny of our data revealed that many kindergartners tried to apply foraging-based facts and adaptation schema to the far generalization camouflage scenarios (50% did so after book 2; 53% did so after a 1-month delay).

It remains for future research to explore whether younger children's generalization performance could have been improved by comparing two more structurally and relationally disparate storybooks (e.g., a foraging and a predation storybook). However, existing research that guided our choice to maintain moderate to high structural similarity across comparison cases (e.g., Loewenstein & Gentner, 2001; Rittle-Johnson & Star, 2007, 2009; see also Vasilyeva & Bowers, 2010) as well as more recently published findings (Shtulman et al., 2016) suggests that improvement on that basis may be unlikely. For example, Shtulman et al. (2016) found that increasing analogical dissimilarity across two brief adaptation training scenarios did not improve young children's learning of individual natural selection principles. It is also possible that kindergartners' far generalization difficulties stemmed from a lack of relevant background knowledge about the concept of camouflage (see Ganea et al., 2011). Introductions to the assessment scenarios used in the present investigation did not state the environmental implications of the focal body parts: Instead, children had to infer that the focal trait was relevant to foraging or hiding from a predator. Performance would have been undermined for children who failed to make this connection. In short, early developing conceptual biases, general processing limitations, limited background biological knowledge, and a more distracting classroom learning setting all may have contributed to kindergartners' more restricted learning and far generalization performance. Further studies are needed to evaluate the potential role of each of these factors in learning about adaption to help inform age-appropriate curriculum.

In marked contrast to their younger counterparts, second graders succeeded not only on near generalization but also far generalization. While they came into the intervention displaying similarly low levels of pretest knowledge as kindergartners, second graders not only gained the isolated facts but also the selectionist logic of natural selection, successfully applying it to both foraging and camouflage adaptation contexts. Although we expected the classroom setting to generally present a more challenging learning environment than the one-on-one setting of prior work, the majority of second graders' gains were displayed immediately following the first storybook reading. At first glance it might then appear that their near ceiling performance following the first book meant they did not benefit from the second part of the intervention; however, there are several reasons to suspect that they did benefit.

To start, extensive prior research demonstrates that comparing examples facilitates learning and generalization abilities (e.g., Bean, Searles, Singer, & Cowen, 1990; Brown & Kane, 1988; Yanowitz, 2001). To succeed on far generalization, children had to possess a conceptual understanding of the mechanism of natural selection that was sufficiently abstract and flexible that it could be applied in a novel predation scenario in which species members with a diminished (rather than amplified) form of the focal trait had a survival advantage (see Ha & Nehm, 2014). Children had to articulate this understanding while also potentially suppressing the foraging-model of adaptation presented in both of the storybooks. Given such challenges, offering children the opportunity to process and unpack the logic of natural selection a second time as well as discuss it likely helped children gain a deeper understanding of the mechanism of adaptation.

Further evidence that the second part of the intervention was beneficial for older children comes from their performance following it. Although not statistically significant, their performance on near generalization following the second book and classroom discussion trended upwards: 100% of children achieved a Level 2 or higher global score compared to 76% at the book 1 near generalization assessment. Closer examination of the data revealed that nearly a quarter of second graders gained a correct theoretical understanding of natural selection after the second phase of the intervention, in part, because they no longer held a misconception about adaptation (see Table 5). Thus, the second part of the intervention seemed to combat older children's incorrect theories about trait variation and species change over time: This held true even in the far generalization contexts in which only one child displayed an overt misconception despite the additional conceptual demands of the assessment.

Second graders' robust abilities to learn and apply the selectionist logic of adaptation has been demonstrated before (Emmons et al., 2016; Kelemen et al., 2014). Their more developed general processing abilities and perhaps greater knowledge about camouflage may have contributed to the successes seen here. However, prior findings also suggest an additional factor: Children at this age display better abilities to entertain the idea of within-species variation than younger children (Emmons et al., 2016; see also Emmons & Kelemen, 2015; Legare et al., 2013). This enhanced capacity would certainly have supported their abilities to represent species' specialized body parts as resulting from changes in trait frequencies within a population over time and successive generations. Future investigations should seek to explore how and when children spontaneously entertain variation and ways to foster and develop this important skill for learning about evolutionary processes.

Conclusion

Findings from this study highlight children's capacities for enduring learning of complex scientific material from a factual narrative picture storybook intervention. They demonstrate that young children participating in urban classrooms can accurately and productively acquire factual and theoretical knowledge about adaptation by natural selection when invited to learn, explain, and compare image-supported explanatory materials constructed according to principles derived from contemporary developmental, education, and learning sciences research. The generalizable mechanistic learning demonstrated here was facilitated by explanatorily rich picture storybooks that causally integrate individual cause-and-effect facts to gradually build a comprehensive and coherent mechanistic explanation free of intentional, teleological, and essentialist allusions. Notably, related work that evaluated children's learning of individual component facts of adaptation and found that many children still had misconceptions about the process utilized much briefer and therefore less causally cohesive narratives than those employed in the present investigation (see Legare et al., 2013; Shtulman et al., 2016).

While we believe the comprehensive nature of the present materials was important to children's learning, at least as important was children's active engagement in explaining and applying their knowledge in the context of the learning assessments. In short, the learning benefits revealed here should not be regarded as simple products of testimony or passive "telling" (Schwartz & Bransford, 1998) but of children's theory-building abilities when offered access to rich cause-and-effect explanatory content that elaborates an overarching mechanism. Given the degree of learning observed here, the present findings run counter to assumptions that counterintuitive mechanistic concepts or abstract principles such as natural selection are beyond the reach of elementary school-aged children. They also yield optimism about the viability of constructing a learning progression on evolution content in elementary school.

More specifically, in considering a learning progression, the present findings challenge conventional educational wisdom that instruction should focus on gradual, piecemeal expertise building. Rather, they support that causally integrated scientific instruction can be introduced early, before biases that are known to impede evolutionary understanding have an opportunity to entrench (see also Emmons et al., 2016; Kelemen, 2012; Kelemen et al., 2014). Given children's natural explanatory drives (Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1989) and natural interest in animals and biology (Kelemen, Callanan, Casler, & Pérez-Granados, 2005; LoBue, Bloom Pickard, Sherman, Axford, & DeLoache, 2013), educators have the potential to make introducing complex evolution content fun and challenging for young learners. The current findings support that children's learning benefits from causally cohesive explanations with sufficient mechanistic detail to prevent explanatory gaps that are subject to reinterpretation. Relatedly, avoiding transformational language suggesting that change occurs at the individual level or is need driven (e.g., "Giraffes evolved long necks to reach food")—and instead using relational language comparing the differences in traits between species members (e.g., "Some giraffes had necks that were longer")—can help to scaffold young children's awareness of within-species variation and understand that change happens at the population level. As the present investigation demonstrates, giving children multiple opportunities to see the same process in novel contexts as well as practice explaining adaptation in analogous scenarios provides children numerous chances to learn and abstract biological factual information along with deeper theoretical explanations that are crucial for long-term, flexible understanding. In sum, child-friendly scientifically accurate storybook materials like the ones employed here are a non-threatening resource for educators that can be used to introduce pivotal content and begin the process of scientific inquiry on complex scientific concepts.

Although these findings demonstrate what is possible with children in early elementary school, future research is needed to explore whether the type of intervention used in the present investigation would be successful with older students and adults. While we would expect the content of the storybook intervention to be accessible to older students and adults, it is currently unknown whether their misconceptions about animals and the origins of their traits would derail their long-term learning (see Gregory, 2009, for a review on misconceptions). Suggestive evidence that adults' learning might be hindered by preexisting misconceptions comes from a museum-based study that explored whether museum visitors' ideas about evolution would change after viewing an evolutionary science exhibit. This study found that while older children, young teens, and adults all gained a better understanding of key evolutionary concepts (e.g., common descent, differential survival), adults alone were more likely to use incorrect intuitive-based reasoning that animals adapt in response to need after the exhibit, despite the exhibit content deliberately avoiding such intuitive language (Spiegel et al., 2012). This finding suggests that adults' belief systems may be less flexible than those of children and, consequently, more strongly guided by scientifically inaccurate assumptions that have been left unchallenged for many years. Additional research is needed to evaluate this possibility in the context of the storybook materials used in this intervention and is currently in progress.

Furthermore, it remains for longitudinal research to see what, if any, successful long-term consequences of initiating learning of counterintuitive concepts like evolution in elementary school might be for science education broadly defined. Under one view of conceptual development, doing so could thwart the development of scientific misconceptions by conceptually changing and replacing early developing explanatory ideas constructed under the influence of intuitive modes of construal. However, based on the dual processing position that guides this research (Emmons et al., 2016; Kelemen et al., 2014), such entire restructuring of early developing, deep-seated, and intuitive ideas may be unlikely. Under this view, the benefit of early learning is not to replace intuitive modes of thinking but rather to build the representational strength of counterintuitive scientific ideas early on such that they compete effectively with scientifically inaccurate intuition under a multitude of reasoning conditions and contexts. Future research will explore what kinds of evolutionary concepts (e.g., speciation) can be accurately learned in early elementary school and what enduring educational outcomes are likely to result.

We would like to thank the families and school that participated in this research. We would also like to thank the Child Cognition Lab Research Assistants, including Connor Gallik, Carolyn Lee, and Laura Jean Nelson, who assisted with testing, designing study materials, and processing data. This research was supported by National Science Foundation Project Grant 1007984 and 50204432 awarded to Deborah Kelemen. The views expressed in this article are those of the authors.

Note

[1] One question that had concluded previous open-ended portions of the assessment was excluded in the present investigation. It probed children's understanding of how long it took for the species to adapt and was not previously found to yield useful or straightforward answers. It was therefore omitted in the interest of time.

References

Achieve, Inc. (2013). Next generation science standards. Retrieved from http://www.nextgenscience.org/searchstandards.

American Association for the Advancement of Science. (2009). Benchmarks for science literacy. New York, NY: Oxford University Press.

Asghar, A., Wiles, J. R., & Alters, B. (2007). Canadian pre-service elementary teachers' conceptions of biological evolution and evolution education. McGill Journal of Education, 42, 189–210.

Atran, S. (1996). Modes of thinking about living kinds. In D. R. Olson, & N. Torrance (Eds.), Modes of thought: explorations in culture and cognition (pp. 216–260). New York, NY: Cambridge University Press.

Bean, T. W., Searles, D., Singer, H., & Cowen, S. (1990). Learning concepts from biology text through pictorial analogies and an analogical study guide. The Journal of Educational Research, 83, 233–237.

Berti, A. E., Barbetta, V., & Toneatti, L. (2017). Third-graders' conceptions about the origin of species before and after instruction: An exploratory study. International Journal of Science and Mathematics Education, 15, 215–232.

Berti, A. E., Toneatti, L., & Rosati, V. (2010). Children's conceptions about the origin of species: A study of Italian children's conceptions with and without instruction. Journal of the Learning Sciences, 19, 506–538.

Bjorklund, D. (2005). Children's thinking: Cognitive development and individual differences. Pacific Grove, CA: Wadsworth.

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

Brown, A. L., Kane, M. J., & Long, C. (1989). Analogical transfer in young children: Analogies as tools for communication and exposition. Applied Cognitive Psychology, 3, 275–293.

Browning, E, & Hohenstein, J (2013). The use of narrative to promote primary school children's understanding of evolution. Education 3–13: International Journal of Primary Elementary and Early Years Education, 43, 530–547.

Carey, S. (1985). Conceptual change in childhood. Cambridge, MA: MIT Press.

Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439–477.

Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: Cognitive principles and the development of biology thinking. CBE-Life Sciences Education, 11, 209–215.

Coley, J. D., & Tanner, K. D. (2015). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. CBE-Life Sciences Education, 14, 1–19.

DeLoache, J. S. (2004). Becoming symbol-minded. Trends in Cognitive Sciences, 8, 66–70.

Dickes, A. C., & Sengupta, P. (2013). Learning natural selection in 4th grade with multi-agent- based computational models. Research in Science Education, 43, 921–953.

Emmons, N. A., & Kelemen, D. A. (2015). Young children's acceptance of within-species variation: Implications for essentialism and teaching evolution. Journal of Experimental Child Psychology, 139, 148–160.

Emmons, N. A., Smith, H., & Kelemen, D. A. (2016). Changing minds with the story of adaptation: Strategies for teaching young children about adaptation. Early Education and Development, 27, 1205–1221.

Evans, E. M. (2000). The emergence of beliefs about the origins of species in school-age children. Merrill-Palmer Quarterly, 46, 221–254.

Evans, E. M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. Cognitive Psychology, 42, 217–266.

Evans, J. S., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. Perspectives on Psychological Science, 8, 223–241.

Ferrari, M., & Chi, M. T. H. (1998). The nature of naive explanations of natural selection. International Journal of Scientific Education, 20, 1231–1256.

Friedman, W. J. (1977). The development of children's understanding of cyclic aspects of time. Child Development, 48, 1593–1599.

Ganea, P. A., Ma, L., & DeLoache, J. S. (2011). Young children's learning and transfer of biological information from picture books to real animals. Child Development, 82, 1421–1433.

Gelman, S. A., & Rhodes, M (2012). Two-thousand years of stasis": How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. K. Brem, E. M. Evans, & G. M. Sinatra (Eds.), Evolution challenges: integrating research and practice in teaching and learning about evolution (pp. 3–21). New York, NY: Oxford University Press.

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

Gentner, D. (1989). The mechanisms of analogical learning. In S Vosniadou, & A Ortony (Eds.), Similarity and analogical reasoning (pp. 199–241). London, England: Cambridge University Press.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. Journal of Educational Psychology, 95, 393–408.

Gopnik, A., & Meltzoff, A. N. (1997). Words, thoughts, and theories. Cambridge, MA: MIT Press.

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

Ha, M., & Nehm, R. H. (2014). Darwin's difficulties and students' struggles with trait loss: Cognitive-historical parallelisms in evolutionary explanation. Science & Education, 23, 1051–1074.

Herrmann, P. A., French, J. A., DeHart, G. B., & Rosengren, K. S. (2013). Essentialist reasoning and knowledge effects on biological reasoning in young children. Merrill-Palmer Quarterly, 59, 198–220.

Horwitz, P., McIntyre, C. A., Lord, T. L., O'Dwyer, L. M., & Staudt, C. (2013). Teaching "evolution readiness" to fourth graders. Evolution: Education and Outreach, 6, 21.

Järnefelt, E., Canfield, C. F., & Kelemen, D. (2015). The divided mind of a disbeliever: Intuitive beliefs about nature as purposefully created among different groups of non-religious adults. Cognition, 140, 72–88.

Jensen, M. S., & Finley, F. N. (1995). Teaching evolution using historical arguments in a conceptual change strategy. Scientific Education, 79, 147–166.

Kahneman, Daniel. (2011). Thinking fast and slow. New York: Farrar, Straus and Giroux.

Keil, F. C. (1989). Concepts, kinds, and conceptual development. Cambridge, MA: MIT Press.

Kelemen, D. (1999). Why are rocks pointy? Children's preference for teleological explanations of the natural world. Developmental Psychology, 35, 1440–1453.

Kelemen, D. (2012). Teleological minds: how natural intuitions about agency and purpose influence learning about evolution. In K. S. Rosengren, S. K. Brem, E. M. Evans, & G. M. Sinatra (Eds.), Evolution challenges: integrating research and practice in teaching and learning about evolution (pp. 66–92). New York, NY: Oxford University Press.

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

Kelemen & The Child Cognition Lab. (2017). How the piloses evolved skinny noses. Boston, MA: Tumblehome Learning.

Kelemen & The Child Cognition Lab. (2018). How the dormits evolved long backs. Boston, MA: Tumblehome Learning.

Kelemen, D., Callanan, M. A., Casler, K., & Pérez-Granados, D. R. (2005). Why things happen: Teleological explanation in parent-child conversations. Developmental Psychology, 41, 251.

Kelemen, D., Emmons, N., Seston-Schillaci, R., & Ganea, P. (2014). Young children can be taught basic natural selection using a picture storybook intervention. Psychological Science, 25, 893–902.

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, 1074–1083.

Legare, C. H., Evans, E. M., Rosengren, K., & Harris, P. (2012). The coexistence of natural and supernatural explanations across cultures and development. Child Development, 83, 779–793.

Legare, C. H., Lane, J., & Evans, E. M. (2013). Anthropomorphizing science: How does it affect the development of evolutionary concepts? Merrill-Palmer Quarterly, 59, 168–197.

LoBue, V., Bloom Pickard, M., Sherman, K., Axford, C., & DeLoache, J. S. (2013). Young children's interest in live animals. British Journal of Developmental Psychology, 31, 57–69.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. Journal of Cognition and Development, 2, 189–219.

Metz, K. E. (1995). Reassessment of developmental constraints on children's science instruction. Review of Educational Research, 65, 93–127.

Metz, K. E. (2009). Rethinking what is "developmentally appropriate" from a learning progression perspective: The power and the challenge. Review of Science, Mathematics and ICT Education, 3, 5–22.

Nadelson, L., Culp, R., Bunn, S., Burkhart, R., Shetlar, R., Nixon, K., & Waldron, J (2009). Teaching evolution concepts to early elementary school students. Evolution Education and Outreach,, 2, 458–473.

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press.

Nehm, R. H., & Schonfeld, I. S. (2007). Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? Journal of Science Teacher Education, 18, 699–723.

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

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. Journal of Educational Psychology, 99, 561–574.

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

Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: Learning outcomes supported by different types of comparisons. In J. P. Mestre, & B. H. Ross (Eds.), Psychology of learning and motivation: cognition in education. San Diego, CA: Elsevier. 55.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. Perspectives on Psychological Science, 1, 181–210.

Rosengren, K. S., Brem, S. K., Evans, E. M., & Sinatra, G. M. (Eds.). (2012). Evolution challenges: Integrating research and practice in teaching and learning about evolution. Oxford, UK: Oxford University Press.

Rosengren, K. S., Gelman, S. A., Kalish, C. W., & McCormick, M. (1991). As time goes by: Children's early understanding of growth in animals. Child Development, 62, 1302–1320.

Rutledge, M. L., & Warden, M. A. (2000). Evolutionary theory, the nature of science and high school biology teachers: Critical relationships. The American Biology Teacher, 62, 23–31.

Samarapungavan, A., & Wiers, R. W. (1997). Children's thoughts on the origin of species: A study of explanatory coherence. Cognitive Science, 21, 147–177.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. Cognition and Instruction, 16, 475–5223.

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

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

Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. Cognitive Science, 32, 1049–1062.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. Cognition, 124, 209–215.

Shtulman, A., Neal, C., & Lindquist, G. (2016). Children's ability to learn evolutionary explanations for biological adaptation. Early Education and Development, 27, 1222–1236.

Siegal, M., & Peterson, C. (1999). Children's understanding of biology and health. New York, NY: Cambridge University Press.

Silvera, D. H., Josephs, R. A., & Giesler, R. B. (2002). Bigger is better: The influence of physical size on aesthetic preference judgments. Journal of Behavioral Decision Making, 15, 189–202.

Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. Evolution: Education and Outreach, 1, 189–195.

Spiegel, A., Evans, E. M., Frazier, B. F., Hazel, A., Tare, M., Gram, W., & Diamond, J. (2012). Changing museum visitors' concepts of evolution. Evolution: Education and Outreach, 5, 43–61.

Tare, M., Chiong, C., Ganea, P., & DeLoache, J. (2010). Less is more: How manipulative features affect children's learning from picture books. Journal of Applied Developmental Psychology, 31(5), 395–400.

To, C., Tenenbaum, H., & Hogh, H. (2016). Secondary school students' reasoning about evolution. Journal of Research in Science Teaching, 54, 247–273.

Vasilyeva, M., & Bowers, E. (2010). Exploring the effects of similarity on mapping spatial relations. Journal of Experimental Child Psychology, 106, 221–239.

Vlaardingerbroek, B., & Roederer, C. J. (1997). Evolution education in Papua New Guinea: trainee teachers' views. Educational Studies, 23, 363–375.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. Annual Review of Psychology, 43, 337–375.

Williams, J. J., Lombrozo, T., & Rehder, B. (2013). The hazards of explanation: overgeneralization in the face of exceptions. Journal of Experimental Psychology: General, 142, 1006–1014.

Yanowitz, K. L. (2001). Using analogies to improve elementary school students' inferential reasoning about scientific concepts. School Science and Mathematics, 101, 133–142.

Supporting Information

Additional Supporting Information may be found in the online version of this article.