Running head: YOUNG CHILDREN CAN BE TAUGHT BASIC NATURAL SELECTION

Young Children Can Be Taught Basic Natural Selection Using A Picture Storybook Intervention

Deborah Kelemen[1], Natalie A. Emmons[1], Rebecca Seston Schillaci[1], and Patricia A. Ganea[2]

[1]Department of Psychology, Boston University, 64 Cummington Mall, Boston, MA 02215

[2]Ontario Institute for Studies in Education, University of Toronto, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada

Psychological Science (in press).

Please address correspondence to:

Deborah Kelemen, Boston University, Department of Psychology, 64 Cummington Mall, Boston, MA 02215, (617) 353-2758, dkelemen@bu.edu

Natalie Emmons, Boston University, Department of Psychology, 64 Cummington Mall, Boston, MA 02215, (617) 358-1738, nemmons@bu.edu

Author Note

Supported by NSF(ROLE) 0529599 to D.K., and NSF(REESE) 1007984 to D.K. and P.A.G. Thanks to James Traniello, Christopher Schneider, Timothy Heeren, Samantha Barry, Kristen Woo, Angel Lillard, Josh Rottman, Hayley Smith.

Abstract

Adaptation by natural selection is a core evolutionary mechanism. It is also one of the most robustly misunderstood scientific processes. Misconceptions are rooted in cognitive biases found in preschoolers, yet concerns about complexity mean that it is generally not comprehensively taught until adolescence. This is long after untutored theoretical misunderstandings are likely to have entrenched. In a novel approach, we explored 5- to 8-year-old's capacities to learn a basic but theoretically coherent mechanistic explanation of adaptation through a custom storybook intervention. Experiment 1 showed that children understood the population-based logic of natural selection and also generalized it. Furthermore, learning endured 3 months later. Experiment 2 replicated these results and showed that children understood and applied an even more nuanced mechanistic causal explanation. Findings demonstrate that, contrary to conventional educational wisdom, basic natural selection is teachable in early childhood. Theory-driven interventions employing picture storybooks with rich explanatory structure are beneficial.

Keywords: evolution, natural selection, learning, children

Adaptation by natural selection is central to understanding the complexity and functional specialization of living things. Despite this, decades of studies have demonstrated that adaptation by natural selection is one of the most widely misunderstood concepts in science. Misconceptions are not only widespread among high school students and undergraduates (Bishop & Anderson, 1990; Brumby, 1984; Nehm & Reilly, 2007; see Gregory, 2009, for review) who are often targets of instruction on the topic but, disturbingly, also among many of the teachers expected to teach them (Nehm, Kim, & Sheppard, 2009; Nehm & Schonfeld, 2007).

The misconceptions held about adaptation are varied. Instead of construing it as a change in trait frequency that occurs because some organisms in a phenotypically variable population survive and reproduce more successfully in an environment over time, students tend to focus on individuals rather than populations as the locus of change. A classic example is the teleological idea that giraffes evolved long necks because they needed to reach high leaves. The error here rests not in believing that trait functionality is relevant to adaptation but instead in the mistaken frameworks of untutored causal assumptions or "intuitive theories" in which that belief is embedded. These include ideas that effortful action on the part of individuals or, alternatively, the personified force of "Evolution" is capable of transforming species members' essential nature so that they attain functionally beneficial, heritable traits (Gregory, 2009). Problematically, these ideas, which focus on goal-directed transformations of individuals within a lifetime rather than the non-goal-directed population-based process of differential survival and reproduction, are resistant to change: Students demonstrate only modest improvements in understanding after sometimes extended instruction on natural selection (Ferrari & Chi, 1998; Jensen & Finley, 1995; Vlaardingerbroek & Roederer, 1997). This learning challenge has broad implications given that natural selection is relevant to understanding not only within-species adaptation—the focus of the current paper—but ultimately also larger scale macro-evolutionary change like speciation.

In terms of understanding the source of the problem, developmental research points in an important direction. From early in development, young children display conceptual biases that, while useful in everyday reasoning, can begin to interact to yield older students' theoretical misconceptions about adaptation (Coley & Tanner, 2012; Rosengren, Brem, Evans, & Sinatra, 2012). For example, children in preschool and early elementary school show: teleological biases to explain the origins of natural object properties by reference to functions (Keil, 1995; Kelemen, 2004); intentionality biases to construe events and objects as intentionally caused (Evans, 2001; Rosset & Rottman, in press); and essentialist biases to view species members as sharing an invariant, inviolable essence (Gelman, 2003; Shtulman & Shultz, 2008). Children are natural explanation seekers who organize their knowledge into theoretical frameworks (Carey, 1985; Gopnik & Meltzoff, 1997; Wellman & Gelman, 1992), and by 6 to 10 years of age, these potentially independent conceptual biases show signs of integrating into intuitive causal theories that connect ideas about biological functionality in nature with notions of invariant essences (Shtulman & Shultz, 2008) and goal-directed design (Kelemen & DiYanni, 2005). In short, a by-product of useful everyday cognition is that the kinds of untutored theories that impede older students' understanding of natural selection are already beginning to coalesce in early elementary school, if not before.

Given these findings, recommended timetables for exposing children to explanations of adaptation are concerning. In the United States, K-12 science education standards suggest that a comprehensive presentation of the logic of adaptation by natural selection occur between grades 8-12 (AAAS, 2009; Achieve, 2013; NRC, 2012). That is, while teaching about some conceptual components of the theory is recommended earlier, instruction explicitly focused on explaining adaptation using a population-based mechanism that comprehensively integrates concepts of within-species variation, environmental context, inheritance, differential survival, and differential reproduction is typically delayed until 13- to 18-years of age (Achieve, 2013, sec. HS-LS4; NRC, 2012, sec. LS4.B; but see Scott, 2012, on uneven evolution standards implementation). The rationale underlying the recommended timing is understandable: Even in its simplest form, adaptation by natural selection is a multi-faceted causally-complex mechanism. It is therefore assumed that children first need gradual tutoring on component isolated facts, such as the connection between food and survival or trait variation and differential survival, before moving to tutoring on the selectionist mechanism as a coherent integrated whole.

However, given children's emerging scientifically inaccurate, untutored theories, it is questionable whether this piecemeal approach to instruction is ideal, especially considering the potential advantages of offering children an age-appropriate but accurate causally-comprehensive version of the theory. The latter alternative not only familiarizes children with the individual facts, it also begins establishing a coherent population-based explanatory framework that, with repeated familiarization, may become habitual enough to resist reinterpretation by biases and competition from typically developing intuitive theoretical ideas. On this view then, an optimal time to begin comprehensively familiarizing children with counter-intuitive scientific explanations is relatively early, during ages when alternative commonsense explanatory frameworks are still relatively fragmentary (e.g., Kelemen & DiYanni, 2005). Furthermore, individual developmental findings suggest that delaying comprehensive instruction on adaptation until adolescence may be unnecessary: By kindergarten, many children already know some isolated biological facts that collectively might support a grasp of the theory. For example, they know that: body parts perform survival functions (Jaakkola & Slaughter, 2002; Keil, 1995; Kelemen, 1999); animals need food to remain healthy and alive (Inagaki & Hatano, 2002); and offspring tend to resemble their birth parents (Gelman & Wellman, 1991; Solomon, Johnson, Zaitchik, & Carey, 1996; Springer & Keil, 1989). Despite having some of these facts, what children do not possess is an alternative to commonsense ways of drawing them together when explaining why animals have functional traits and show signs of apparent design. In this research, we therefore sought to capitalize on their natural theory-building drives to offer them one.

Leveraging findings on young children's biological factual knowledge (see Gripshover & Markman, 2013), their natural interest in trait function, and the likely fragility of emerging intuitive theories, in two experiments, we explored 5- to 8-year-olds' abilities to understand and apply a basic but comprehensive explanation of within-species adaptation by natural selection through a carefully crafted custom picture storybook intervention. We used a picture storybook because the format is child friendly, invites a beneficial joint-attentional learning context and the image-enriched narrative reduces cognitive load (Mayer & Moreno, 2003) while supporting a multi-faceted causal explanation (see Brown, Kane, & Long, 1989; Browning & Hohenstein, 2013; Legare, Lane, & Evans, 2013, for other narrative-based approaches with related but different goals). Additionally, young children have been found to learn and generalize simple biological facts from picture books to real animals (Ganea, Ma, & DeLoache, 2011).

Despite theoretical reasons for targeting early elementary ages, young children's information-processing limitations (Bjorklund, 2005; Friedman, 1977) nevertheless gave us reasons to suspect even a basic version of the logic of adaptation would be too hard. In Experiment 1, we therefore began with a storybook describing a more easily conceptualized case: rapid natural selection in a fictional mammalian population ("pilosas") experiencing sudden die-off due to the effects of extreme climate change on the location of their food source. The narrative focused on the immediate impact on the pilosa population of their insect food moving underground into deep, narrow tunnels. Each page of the narrative incorporated a new fact that mechanistically elaborated how differential survival and reproduction caused the pilosas to go from having substantial variation in trunk size (wide and thin trunks) to less variation (thin trunks predominating). In addition to a pre-test assessment, children's comprehension and generalization of the storybook explanation was evaluated with two assessments immediately after storybook reading and two more again three months later. Based on Experiment 1 results, Experiment 2 explored children's comprehension and generalization of an even more nuanced explanation of adaptation: Rather than focusing on the initial population and their immediate offspring, in Experiment 2, the storybook emphasized gradual natural selection over multiple generations.

Experiment 1

Method

Participants. Twenty-eight 5- to 6-year-olds (17 males; *M* = 5;9, *SD* = 6 months) and thirty-three 7- to 8-year-olds (15 males; *M* = 7;9, *SD* = 5 months) were recruited from Boston (73% Caucasian, 10% Asian, 2% Hispanic, 2% African-American, 13% other). A subset (younger: *n*= 21; older: *n*= 23) returned three months later for Day 2 testing. Parent questionnaires indicated that children came from backgrounds without marked natural selection knowledge. Children were tested individually for 60 minutes on Day 1 and 30 minutes on Day 2.

Materials and procedure. The custom-written ten-page storybook employed realistic pictures and factual narrative using non-teleological, non-intentional language to answer the question posed at the book's beginning: why did pilosas go from having highly variable trunk width in the past to having predominantly thin trunks nowadays? The explanation then unfolded, tightly causally connecting information on six natural selection concepts: trait variation within a population, habitat and food source change in response to abrupt climate change, differential health and survival due to differential food access, differential reproduction due to differential health, trait inheritance, and trait frequency change over multiple generations. While multiple generations were depicted, most of the book focused on describing adaptation in the initial population and their immediate offspring. Reading took 10 minutes.

Children's understanding of basic natural selection was assessed with a novel animal population before storybook exposure (Pre-test Day 1) and twice immediately afterwards: once to explore children's comprehension of the population-based logic of the pilosa storybook (Comprehension Day 1) and once to explore their ability to generalize it to a novel species (Generalization Day 1). Long-term retention was explored with a subset of children 3 months later via a second comprehension assessment about the pilosas (Comprehension Day 2) and a second generalization assessment with yet another novel species (Generalization Day 2). Each conceptually parallel assessment comprised: (a) five closed-ended questions with requests for answer justifications that evaluated children's knowledge of component isolated facts relevant to the natural selection explanation (e.g., the relationship between: food and health; health and fecundity); and (b) five open-ended questions probing children's capacity to self-generate a causally-coherent explanation of adaptation that integrated knowledge of the component isolated facts. The most central of these questions straightforwardly asked children to explain the change in trait frequency across time (i.e., why do pilosas only have thin trunks now?). Self-generating accurate explanations after storybook exposure was presumed to facilitate children's comprehension and abstraction of the causal logic. Importantly, however, children never received corrective feedback: Children who failed to grasp the causal logic were therefore likely to falter across all post-test assessments. Furthermore, open-ended questions and follow-up prompts were structured so that they would elicit children's underlying inaccurate causal ideas (e.g., transformationist misconceptions) as well as their accurate ones. Tables S1, S2, and S3 in Supplemental Materials available online provide all questions used in Experiments 1 and 2 with sample responses.

Each assessment began by introducing children to the fictional species under question via four realistic pictures showing: the ancestral population, the ancestral habitat, the contemporary population, and the contemporary habitat. Children then received the standard set of 10 assessment questions. Children answered closed-ended questions about isolated facts by pointing between picture pairs illustrating alternative answers and justifying their responses. Open-ended questions were accompanied by pictures of the ancestral and contemporary populations that children could reference when explaining why the species changed over time and what happened to physically disadvantaged and advantaged members. The species presented in pre-test, comprehension, and generalization assessments were physically dissimilar to each other (e.g., birds, okapi-like mammals) and had unique habitats. In light of numerous disparities in surface structure that resulted from using dissimilar species and environmental contexts in each assessment, a focus on explaining adaptation of traits somehow related to food-acquisition (e.g., necks, trunks, beaks) held across all assessments. This was because generalization is recognized as one of the hardest tasks in education and prior research indicated (e.g., Gentner, 1989) that we were already substantially challenging children's transfer abilities with the variabilities in surface structure already introduced.

A conceptual checklist and conservative coding rubric that weighted children's open-ended causal explanation was applied to each assessment. Overall understanding of natural selection was then categorized into one of five hierarchical levels for each assessment. Supplemental Materials online provide coding details. In Level 0, "No isolated facts," children's responses to closed-ended questions demonstrated insufficient knowledge of the requisite isolated facts to be credited with any understanding of natural selection (< 4 out of 5 closed-ended questions correct). In Level 1, "Isolated facts but no natural selection understanding," responses to closed-ended questions revealed sufficient knowledge of isolated facts (≥ 4 out of 5 closed-ended questions correct) but an inability to integrate those facts into a coherent accurate self-generated explanation of population-based change absent misconceptions. In Level 2, "Foundation for natural selection understanding," closed-ended responses demonstrated sufficient isolated factual knowledge plus an accurate, causally-coherent yet incomplete self-generated population-based explanation focused on adaptations arising through differential survival. In Level 3, "Natural selection understanding in one generation," responses revealed sufficient factual knowledge and an accurate, self-generated population-based explanation that adaptations arise through differential survival and differential reproduction but children limited their focus to the initial population and their immediate descendants. Level 4, "Natural selection understanding in multiple generations," was similar to Level 3, but self-generated explanations also referenced that natural selection occurs over multiple generations. To underscore, in contrast to other explorations of children's evolutionary ideas (e.g., Browning & Hohenstein, 2013; Legare et al., 2013), children in this study were only ever credited with any understanding of natural selection (Level 2 or higher) when there were no signs of transformationist misconceptions that individuals acquire advantageous traits within their lifetime. Inter-rater reliability between two coders was excellent (Kappa = .84).

Results

Younger children. Treating Pre-test Day 1 as the baseline, analyses were performed to examine how the distribution of children across the 5 hierarchical levels of natural selection understanding changed after storybook exposure. Repeated measures ordinal logistic regressions comparing younger children's levels of natural selection understanding on each of the five assessments revealed that the intervention induced learning, Wald $\chi^2(4) = 33.29$, $p < 0.001$ (see Figure 1). Odds ratios from this analysis indicated the magnitude of change in the odds that children's understanding of natural selection would go up one or many levels between assessment times. Specifically, given their starting levels of understanding at Pre-test Day 1, children's odds of being in a higher level of natural selection understanding at Comprehension Day 1 increased eighteen fold, OR = 18.68, $p < 0.001$, 95% CI [6.74, 51.73]: At Pre-test Day 1, 82% of children were at Level 0, displaying insufficient knowledge of the isolated facts to support natural selection understanding. This dropped to 11% after storybook exposure. This change was not simply due to children acquiring an atheoretical understanding of isolated facts. Before hearing the story, only 11% of children displayed a population-based logic. After hearing the story, 54% had integrated the facts into an accurate population-based explanation, incorporating at minimum the concept of differential survival (Level 2 and higher). In addition to being able to understand the population-level logic of the storybook, children successfully generalized it to an entirely new animal despite the challenges of transfer (Brown et al., 1989): There was no significant change in children's odds of being in a higher level of natural selection understanding between Comprehension Day 1 and Generalization Day 1, $p = 0.14$. Younger children's learning also endured: They showed no significant change in odds between Comprehension Day 1 to either Comprehension Day 2, $p = 0.06$, or the more challenging assessment of Generalization Day 2, $p = 0.39$, three months later.

Older children. The intervention also induced learning among older children, Wald $\chi^2(4) = 31.51$, $p < 0.001$ (see Figure 1). Many older children entered the experiment already possessing sufficient knowledge of the isolated facts and even some theory. Nevertheless, children's odds of being in a higher level of natural selection understanding increased eleven fold between Pre-test Day 1 and Comprehension Day 1, OR = 11.54, $p < 0.001$, 95% CI [4.78, 27.86] because the storybook intervention bolstered their factual knowledge and ability to integrate those facts into a coherent population-based theory. After hearing the storybook, the proportion of children with sufficient knowledge of the isolated facts increased from 57% to 93%, with 90% displaying a Level 2 or higher understanding of natural selection at Comprehension Day 1. While only 9% of children displayed a Level 3 or 4 understanding of natural selection at Pre-test Day 1, this rose to 48% at Comprehension Day 1. Although there was a small two-fold decrease in children's odds of being in a higher level of natural selection understanding between Comprehension Day 1 and Generalization Day 1, OR = 0.47, $p = 0.03$, 95% CI [0.24, 0.91], children were largely successful in applying what they learned from the storybook to a novel animal: 79% continued to display a Level 2 or higher understanding of natural selection. This small drop in performance disappeared when children were assessed again three months later. Children showed no change in their odds of being in a higher level of natural selection understanding between Comprehension Day 1 and Comprehension Day 2, $p = 0.14$, or between Comprehension Day 2 and Generalization Day 2, $p = 0.22$. As with younger children, older children's learning was therefore not only robust and generalizable but also endured over time.

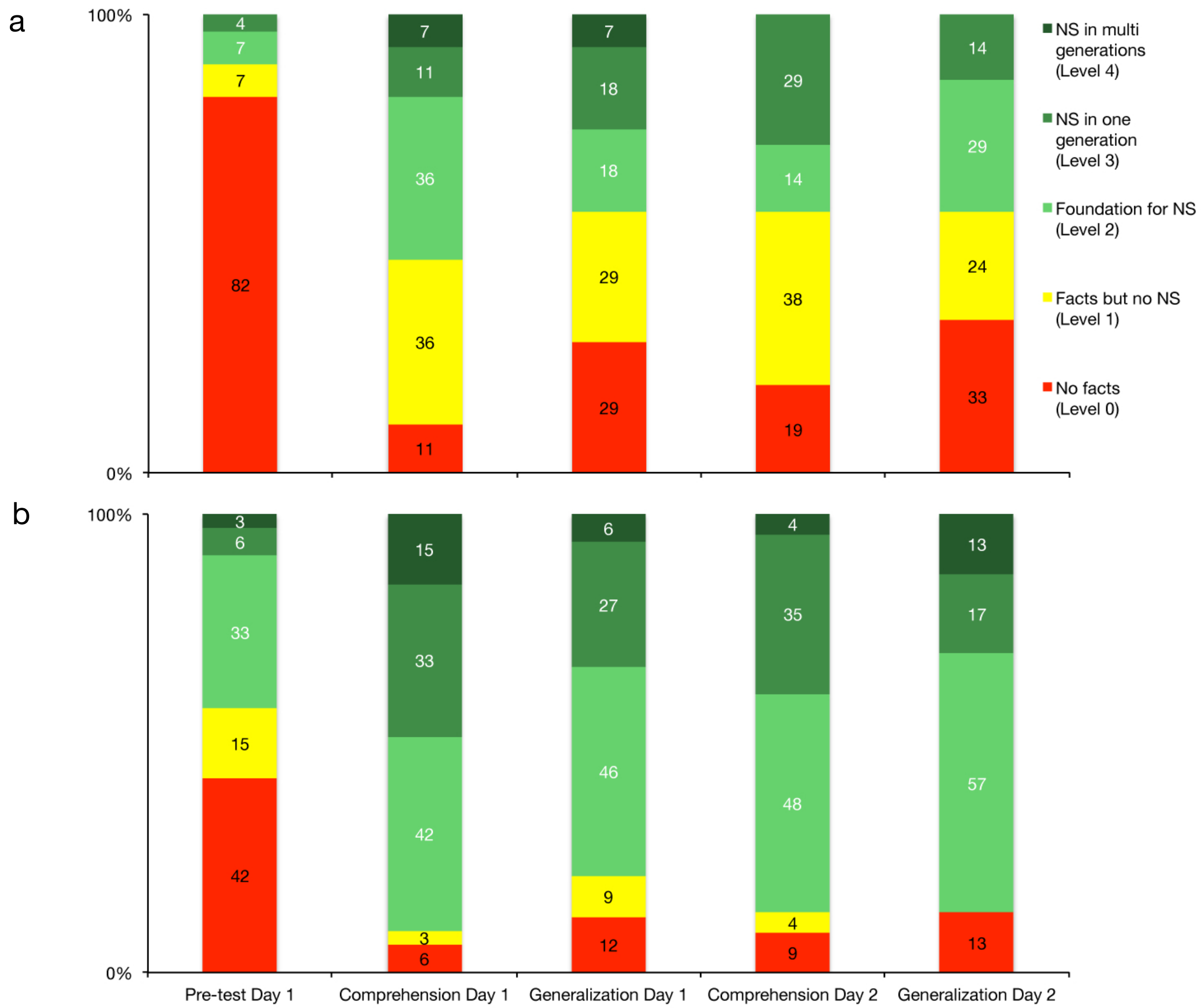

Fig. 1. Percentages of (a) younger and (b) older children classified into the five levels of natural selection (NS) understanding on the five assessments in Experiment 1.

Discussion

Experiment 1 provided initial evidence that, contrary to conventional educational wisdom, young children can grasp the population-based logic of natural selection when it is presented in a basic, cohesive, comprehensive way: Five- to eight-year-olds showed substantial learning from hearing and talking about the 10-page storybook. Furthermore, their understanding was coherent in nature. Children not only demonstrated increased knowledge of isolated biological facts but also integrated them into a cogent population-based understanding of adaptation when self-generating explanations to open-ended questions that pushed them to reveal the accuracy of their underlying reasoning. Despite the absence of corrective feedback, this understanding was then transferred to new cases and retained over time, with children's levels of understanding remaining constant over three months. Comprehension and the challenging task of generalization were particularly pronounced among 7- to 8-year-olds. Transcripts suggested this was due to their enhanced verbal and information-processing skills.

Such results offered substantial reasons for optimism about children's explanatory capabilities and the instructional format represented by the storybook intervention. However, the unanticipated degree of learning raised questions about children's potential for even greater mechanistic sophistication. Because of concerns about children's information-processing limitations, including their abilities to represent extended time (e.g., Friedman, 1977), Experiment 1's storybook presented children with a case of rapid natural selection wherein adaptation largely occurred because of differential survival and reproduction in the first generation of pilosas born after the weather changed. Perhaps unsurprisingly, many children focused their explanations on the initial generation too. In Experiment 2, we therefore modified the storybook to present a more gradual process emphasizing differential reproduction over multiple generations. This allowed us to explore children's ability to understand a more nuanced, complex explanation of adaptation and the replicability of Experiment 1.

Experiment 2

Method

Participants. Sixteen 5- to 6-year-olds (10 males; M = 6;0, SD = 4 months) and sixteen 7- to 8-year-olds (7 males; M = 8;3, SD = 3 months) were recruited from Boston (75%

Caucasian, 6% Asian, 3% Hispanic, 6% African-American, and 9% other race). Testing took about 60 minutes. Children came from backgrounds without marked natural selection knowledge.

Materials and procedure. Experiment 2 had the same design as Experiment 1 but focused on immediate comprehension and generalization: Children performed a pre-test with a novel species (Pre-test), a comprehension assessment on the pilosas (Comprehension), and a generalization assessment involving another novel species (Generalization) on one day. Three-month delayed assessment was not possible due to high participant attrition over summer vacation.

The revised storybook causally connected the same six concepts as the earlier version. Additionally, the book explicitly incorporated the concept of trait constancy to highlight that the kind of inherited trait an offspring displays at birth does not change kind during an individual's lifetime in response to need. To emphasize a gradualist process of natural selection, disadvantaged pilosas no longer experienced immediate die-off when the climate and location of their food changed. Instead, the number of animals inheriting the more disadvantaged trait diminished over time due to gradual differential reproduction. Images visually represented the numerical takeover of reproductively successful pilosas over successive generations.

Experiment 2 assessments were structured as in Experiment 1 but involved six closed-ended isolated fact questions and four open-ended questions exploring children's capacities to self-generate the logic of natural selection. Compared to Experiment 1, children received additional prompts when self-generating their explanations (e.g., "what happened next?") to further elicit their reasoning in all assessments. Because prompts focused on encouraging children to elaborate their own prior utterances, they had two facets: They could more clearly reveal misconceptions underlying an abbreviated apparently accurate initial open-ended response or reveal greater mechanistic understanding than initial utterances implied (see examples in Supplemental Materials). Finally, pragmatics that potentially caused older children's mild performance dip between Comprehension Day 1 and Generalization Day 1 in Experiment 1 were addressed: In Experiment 2, one experimenter performed the Pre-test, storybook reading, and Comprehension assessment, but another conducted the Generalization assessment to counteract children abbreviating their answers to avoid redundantly repeating themselves to one person. Inter-rater reliability between two coders was excellent (Kappa = .89).

Results

Younger children. Repeated measures ordinal logistic regressions examined how the distribution of children across the five hierarchical levels of natural selection understanding changed after storybook exposure. These revealed that the revised storybook induced learning, Wald $\chi^2(2) = 25.25$, $p < 0.001$ (see Figure 2). Given their levels of understanding at Pre-test, children's odds of being in a higher level of natural selection understanding at Comprehension increased a substantial forty-two fold, OR = 42.17, $p < 0.001$, 95% CI [9.73, 182.78]: At Pre-test, 69% were at Level 0 and no child displayed a population-based grasp of natural selection. After hearing the storybook, only 13% of children lacked the isolated facts and 82% displayed a Level 2 or higher population-based understanding. Indeed, 69% of children incorporated differential reproduction into their explanations to reach Levels 3 or 4. At Generalization, 51% of children continued to describe a population-based mechanism. Even with these impressive gains, there was, however, a small four-fold decrease in children's odds of being in a higher level of natural selection understanding between Comprehension and Generalization, OR = 0.27, $p = 0.01$, 95% CI [0.11, 0.71].

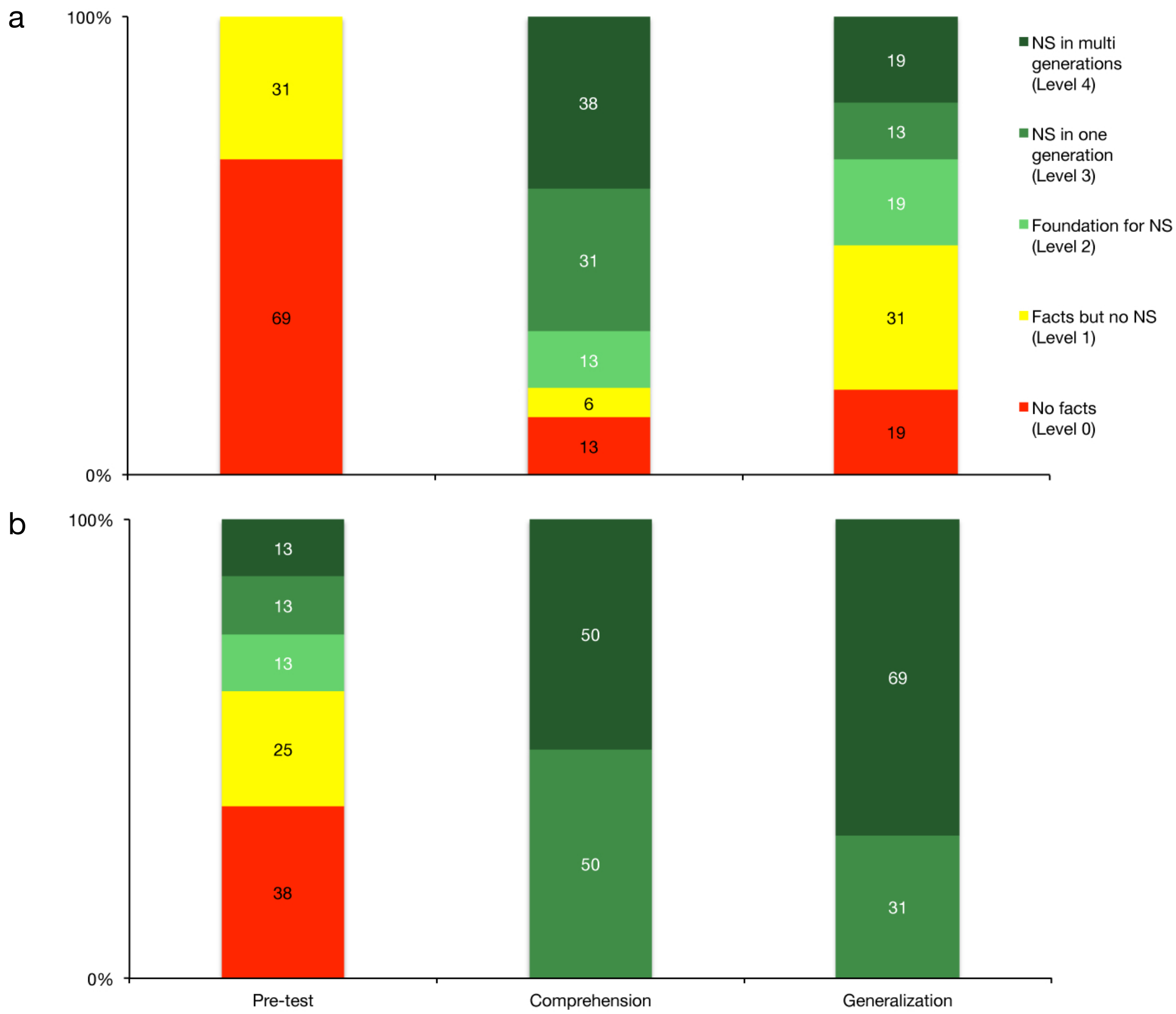

Fig. 2. Percentages of (a) younger and (b) older children classified into the five levels of natural selection (NS) understanding on the three assessments in Experiment 2.

Older children. The revised intervention also induced learning in older children, Wald $\chi^2(2) = 16.72$, $p < 0.001$ (see Figure 2). Older children's odds of being in a higher level of natural selection understanding increased thirty-nine fold from Pre-test to Comprehension, OR = 38.98, $p < 0.001$, 95% CI [5.64, 269.67]. Among older children, 63% were at Level 0 or 1 with no population-based explanation. After the storybook, this dropped to 0% because 100% of children incorporated differential survival and reproduction into their description of adaptation. Fifty percent displayed the highest level of understanding (Level 4), describing natural selection in multiple generations. Children successfully applied what they learned to a novel animal, demonstrating no change in their odds of being in a higher level of natural selection understanding from Comprehension to Generalization, $p = 0.19$.

Discussion

Experiment 2 replicated and extended Experiment 1 findings. Results confirm that early elementary school-aged children can be taught the basic logic of adaptation by natural selection via a brief but comprehensive storybook intervention. Furthermore, the logic that children can grasp is relatively nuanced. Both younger and older children showed abilities to understand that adaptation involves an extended process combining differential survival and reproduction. Older children, particularly, showed substantial capacities to generalize the explanation to novel animals. Indeed, the more detailed theoretical explanation in the second storybook appeared to help older children learn the process of adaptation.

General Discussion

Current findings reveal that, despite its complexity, the basic population-based logic of natural selection is within elementary school-aged children's reach. Young children demonstrated substantial learning of within-species adaptation on the basis of a brief but comprehensive, theoretically-motivated storybook intervention. Gains were particularly marked in Experiment 2 where the intervention resulted in approximately 40-fold increases in children's odds of improving their theoretical understanding. Moreover, in both experiments, children generalized to novel cases despite the known difficulties of transfer. Both age groups learned a great deal, but consistent with their enhanced linguistic and processing capacities, 7- to 8-year-olds showed especially robust abilities to suppress any emergent competing commonsense ideas and master task demands to abstract and transfer the mechanism to markedly different species.

The present results suggest that comprehensive instruction about core evolutionary mechanisms can begin earlier than is currently recommended. Consistent with views of children as natural theory-builders, young children showed remarkable capacities to comprehend and abstract not only isolated facts but mechanistically rich novel scientific explanations when presented in a cohesive framework. Indeed they profited from mechanistic detail: Even children whose performance and knowledge of relevant individual facts was weak at pre-test learned much transferable knowledge from the storybook intervention.

Collectively, such findings offer reasons for optimism regarding effective ways to foster accurate, generalizable basic natural selection understanding. They suggest that leveraging young children's drive for coherent explanation, factual knowledge, interest in trait function, and picture storybooks is a viable initial step towards overcoming conceptual pitfalls that can undermine later learning about adaptation. In concluding this, however, several qualifications must be made. First, although the carefully designed intervention used here yielded substantial learning benefits, it represents the beginning, not the end, of a learning process: This investigation focused on young children's capacities to accurately causally connect the essential components of within-species adaptation by natural selection without misconceptions. Despite the key relevance of this basic mechanism to understanding larger scale evolutionary changes, teaching adult-level detail and promoting children's understanding or acceptance of speciation or common descent was not our goal. As such, this intervention should not be misconstrued as a panacea to all challenges faced by educators teaching a range of evolutionary concepts to older students (Rosengren et al., 2012). Nevertheless, these findings represent a promising first step. Repeated, spaced instruction on gradually scaled-up versions of the logic could ultimately place students in a better position to suppress competing intuitive theoretical explanations such that they can elaborate a richer, more abstract, and broadly applicable knowledge of natural selection. Storybook interventions like the ones reported here seem a promising start from which to foster scientific literacy longer-term.

References

Achieve, Inc. on behalf of the twenty-six states and partners that collaborated on the NGSS (2013). *Next Generation Science Standards.* Retrieved from http://www.nextgenscience.org/

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

Bishop, B.A., & Anderson, C.W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching, 27*, 415-427.

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

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*. doi: 10.1080/03004279.2013.837943

Brumby, M.N. (1984). Misconceptions about the concept of natural selection by medical biology students. *Science Education, 68*, 493-503.

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

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

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

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., (2003). *The essential child: Origins of essentialism in everyday thought.* New York: Oxford University Press.

Gelman, S.A., & Wellman, H.M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition, 38*, 213-244.

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

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.

Gripshover, S.J., & Markman, E.M. (2013). Teaching young children a theory of nutrition: Conceptual change and the potential for increased vegetable consumption. *Psychological Science, 24*, 1541-1553.

Inagaki, K., & Hatano, G. (2002). *Young children's naive thinking about the biological world*. New York, NY: Psychology Press.

Jaakkola, R.O., & Slaughter, V. (2002). Children's body knowledge: Understanding 'life' as a biological goal. *British Journal of Developmental Psychology, 20*, 325-342.

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

Keil, F.C. (1995). The growth of causal understandings of natural kinds. In D. Sperber, D. Premack, & A.J. Premack (Eds.), *Causal cognition: A multi-disciplinary debate* (pp. 234–262). Oxford, England: Clarendon.

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

Kelemen, D. (2004). Are children "intuitive theists"?: Reasoning about purpose and design in nature. *Psychological Science*, *15*, 295-301.

Kelemen, D., & DiYanni, C. (2005). Intuitions about origins: Purpose and intelligent design in children's reasoning about nature. *Journal of Cognitive Development, 6,* 3-31.

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.

Mayer, R. & Moreno, R. (2003). Nine ways to reduce cognitive load in multi-media learning. *Educational Psychologist, 38*, 43-52.

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

Nehm, R.H., Kim, S.Y., & Sheppard K. (2009). Academic preparation in biology and advocacy for teaching evolution: Biology versus non-biology teachers. *Scientific Education, 93*, 1122-1146.

Nehm, R.H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *Bioscience, 57*, 263-272.

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

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.* New York, NY: Oxford University Press.

Rosset, E., & Rottman, J. (in press). The big "whoops!" in the study of intentional behavior: An appeal for a new framework in understanding human actions. *Journal of Cognition and Culture.*

Scott, E. (2012). Foreword. 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. ix-xiii). New York, NY: Oxford University Press.

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

Solomon, G.E.A., Johnson, S.C., Zaitchik, D., & Carey, S. (1996). Like father, like son: Young children's understanding of how and why offspring resemble their parents. *Child Development*, *67,* 151-171.

Springer, K., & Keil, F.C. (1989). On the development of biologically specific beliefs: The case of inheritance. *Child Development*, *60,* 637-648.

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.

Coding Procedure

Using transcriptions of video recordings, coders remained blind to participant age and whether an assessment was a pre-test or generalization assessment (counterbalanced between subjects). Each assessment was assigned one of five overall levels of natural selection understanding (Levels 0-4). Levels were determined using a conceptual checklist and conservative coding rubric that considered all closed-ended and open-ended responses on a given assessment (see Table S3): Level 0, “No isolated facts,” was assigned when children failed to demonstrate sufficient factual knowledge assessed by the closed-ended questions. Level 1, “Isolated facts but no natural selection understanding,” was assigned when children demonstrated sufficient knowledge of isolated facts but no accurate population-based theory of natural selection. This occurred if children failed to correctly connect relevant conceptual components in their open-ended responses or if they demonstrated an active misconception (e.g., claiming individuals acquired advantageous traits). Levels 2, 3, and 4 were assigned when children demonstrated sufficient factual knowledge and an accurate population-based mechanism; however, the three levels differed in the degree of sophistication of the population-based logic. Level 2, “Foundation for natural selection understanding,” was assigned when open-ended responses accurately described adaptation occurring as a result of differential survival due to differential access to food; Level 3, “Natural selection understanding in one generation,” was assigned when children causally connected differential survival and differential reproduction in their open-ended responses to explain adaptation but limited their discussion to one generation; and Level 4, “Natural selection understanding in multiple generations,” was assigned when children extended a Level 3 understanding to include a discussion of differential reproduction occurring over multiple generations.

Coding Details. As Table S3 shows, children had to display sufficient knowledge of isolated facts relevant to natural selection to potentially be credited with any natural selection understanding.

Credit for each isolated fact required children to chose a correct closed-ended answer option and correctly justify their choice. “I don’t know” was coded as inaccurate (see Table S1 for examples).

Open-ended questions probed children’s abilities to self-generate a causally-coherent population-based explanation of why the species changed over time. Self-generated responses were coded for causal reference to three processes: differential survival, differential reproduction in one generation, and differential reproduction in multiple generations. Credit for understanding differential survival was given if children correctly integrated health information with information about differential access to food (e.g., “the ones with wide trunks died because they couldn’t reach the food”). Credit for understanding differential reproduction was given if children either mentioned that animals with advantageous traits had more babies than those with disadvantageous traits (e.g., “the thinner trunks were healthy enough to have babies”) or that animals with disadvantageous traits had fewer babies than those with advantageous traits. Suggestions that animals with disadvantageous traits were equally or more healthy than animals with advantageous traits or that disadvantaged animals were equally or more fecund than animals with advantaged ones were coded “inaccurate.” Because the intervention never used them, if children mentioned terms like “evolve” or “adapt” when responding, they were prompted to explain the meaning. Credit for understanding that natural selection occurs via differential reproduction over multiple generations was given if children either mentioned that babies of animals with advantageous traits would grow up to have babies (e.g., “their children had children”) or that babies of animals with disadvantageous traits would grow up to have no or few babies. Because no assessment questions directly probed children’s awareness of natural selection occurring over multiple generations, children were given credit for this concept if it was mentioned during any part of the assessment. Any reference to ideas demonstrating incorrect transformationist theories that individual members of a population acquired advantageous traits within their lifetimes were coded as misconceptions. These included suggestions that individual animals acquired traits via development (e.g., "when they were a little older they could have some thinner trunks"), ingesting food (e.g., "[they got bigger] because they ate so much"), or functional need (e.g., "[the wider trunks changed because] they needed thinner trunks to reach the food"). Children displaying any misconception were automatically assigned to Level 0 or 1, thus receiving zero credit for understanding natural selection.

This conservative coding scheme was enabled by an important feature of the design: In both experiments, the critical open-ended question asking children to explain species change was followed by follow-up questions (Experiments 1 and 2) and systematic prompts (Experiment 2) encouraged children to elaborate their underlying reasoning. This elicitation approach was adopted because participants were young and unsurprisingly reticent when asked challenging, complex questions: their abbreviated initial responses could mask misconceptions (and conversely, competence). A Level 1 generalization assessment sample response from Experiment 2 highlights these points: Through prompting, the child reveals a misconception not unambiguously apparent in an initial open-ended response even as he clearly incorporates factual elements from the storybook. Note that prompting involved asking "why" and repeating back statements already issued by the child. Leading was therefore avoided because experimenters never added new information.

Experimenter: Many hundreds of years ago most of the grown-up Wilkies had shorter legs but now most of the grown-up Wilkies have longer legs. How do you think that happened?

Child: Because they evolved with..um..longer legs because that's what they needed to be able to survive (*potential misconception*).

Experimenter: When you say evolve, what do you mean?

Child: Evolve means, um, turn into.

Experimenter: Turn into?

Child: Yeah, they turn into…all these wilkies turn into, um, ones with longer legs. (*misconception*)

Experimenter: What happened to wilkies with shorter legs?

Child: They died.

Experimenter: Why?

Child: Because, um, because they couldn't reach the yellow berries.

Experimenter: What happened to the wilkies with the longer legs?

Child: They lived a happy life because they could reach the berries.

Experimenter: Why?

Child: Because they had long legs so they could reach up.

Experimenter: What happened next after they lived a happy life and could reach the berries?

Child: They had kids and it went on and on and on and on and on and on and on… (shortened for length).

e S1. Closed-ended isolated fact questions for Experiment 1 and Experiment 2 with sample justifications.

	Experiment 1			**Experiment 2**		
cept	**Question**	**Accurate Justification**	**Inaccurate Justification**	**Question**	**Accurate Justification**	**Inaccurate Justification**
erential vival	After the weather changed, which group of *okapis [long or short necks]* got more food? Why?	Long necks, because they can reach higher.	Long necks, because they had more room.	Nowadays, will a *wilkie with shorter legs* probably be healthy and live for a long time? Why?	No, because the berries got higher and they couldn't reach it.	No, because they are older.
	After the weather changed, which group of *passerines [big or small beaks]* were less healthy? Why?	Small beaks, because they got less food.	Small beaks, because there was no sun.	Nowadays, will a *rudoo with a longer neck* probably be healthy and live for a long time? Why?	Yes, because the red fruit are up on the top of the trees and it has a long neck.	I don't know.
erential roduction	After the weather changed, which group of *pilosas [thin or wide trunks]* had more babies? Why?	Thin trunks, because they are more healthy.	Thin trunks, because they just got the babies.	Nowadays, will a *rudoo with a shorter neck* probably be healthy and live for a long time? Why?	No, because it had shorter necks so it didn't have enough to eat.	No, because it doesn't have room for the babies to fit in.
	When these baby *hemmies* grow up, which one *[long or short beak]* is more likely to have a baby? Why?	Long beak, because they are more healthy.	Long beaks, because all the other beaks will have the same beak as it.	Nowadays, will a *wilkie with longer legs* probably have lots of children? Why?	Yes, because they're healthy 'cause they eat the fruit from the trees.	Yes, because the appetite is way better because of the legs.
t wledge	See this *okapi* with a *short neck*? If this *okapi* had a baby, what kind of *neck [long or short]* would the baby have? Why?	Short neck, because usually the mother has the same thing as the baby.	Long neck, because they have to eat and they use their long neck.	These grown-up *wilkies* both have *shorter legs*. If these two *wilkies with shorter legs* had a child, what kind of *legs [longer or shorter]* would their child probably have? Why?	Shorter legs. Because the wilkie's parents had shorter legs.	Shorter legs. Because it's just a little child.
				See this young *rudoo*. It was born with *a longer neck*. When this *rudoo* grows up to be an adult, what kind of *neck* will it have *[longer or shorter]*?	Shorter neck. Because it already had a shorter neck when it was born so it should have a shorter neck when it's older.	Longer neck. When that one grows up, it would have to have a long neck to be able to survive.

Note. Italicized information differed depending on the animal species under consideration.

Table S2. Open-ended questions for Experiment 1 and Experiment 2.

Experiment 1	**Experiment 2**
Pilosas had all different sized *trunks* a long time ago, but now *pilosas* only have *thin trunks,* why do you think that happened?	Many hundreds of years ago most of the grown-up *pilosas* had *wider trunks* but now most of the grown-up *pilosas* have *thinner trunks*. How do you think that happened?
What happened to *pilosas* with *thin trunks*?	What happened to *pilosas* with *thinner trunks*? Why? What happened next after…? [repeat child's response to previous question] Why? What happened next after…? [repeat child's response to previous question] Why?
What happened to *pilosas* with *wide trunks*?	What happened to *pilosas* with *wider trunks*? Why? What happened next after…? [repeat child's response to previous question] Why? What happened next after…? [repeat child's response to previous question] Why?
Hundreds of years after the weather changed, were there any families with *thin trunks* in the group? Why?	Did it take a short time or a long time for *pilosas* to go from having mostly *wider trunks* in the past to having mostly *thinner trunks* now? Why?
Hundreds of years after the weather changed, were there any families with *wide trunks* in the group? Why?	

Note. Italicized information differed depending on the animal species under consideration. Questions were in fixed order.

Table S3. Conceptual checklist of natural selection (NS) understanding with examples of open-ended responses in Experiments 1 and 2.

Level	Overall Category	Checklist	Open-ended Response Example
0	No isolated facts	Lacks sufficient knowledge of isolated facts[1]	N/A
1	Isolated facts but no NS understanding	Has sufficient knowledge of isolated facts[1] but one, or more, of the following are also present: - Misconception - No mention of differential survival advantage - Inaccurate mention of any of the three key conceptual components: differential survival advantage, differential reproduction in one generation, differential reproduction in multiple generations	*Level 1 response: Misconception*[2] E: …now pilosas only have thin trunks. Why do you think that happened? P: All the wide trunks became small trunks so they could go into the holes. E: What happened to the pilosas with thin trunks? P: They just stayed the same and they kept eating E: What happened to the pilosas with wide trunks? P: They couldn't eat for a long time so they just waited until their trunks were small. *Level 1 response: No mention of differential survival*[2] E: …now passerines only have big beaks Why do you think that happened? P: They have small beaks and big beaks and it started to rain and the sun came out. E: What happened to the passerines with big beaks? P: They were scared of the rain. E: What happened to the passerines with small beaks? P: They don't cry.
2	Foundational NS understanding	All of the following are present: - Sufficient knowledge of isolated facts - No misconception - Accurate mention of differential survival advantage	*Level 2 response: Differential survival, no differential reproduction*[2] E: …now pilosas only have thin trunks. Why do you think that happened? P: The wide trunks couldn't fit underground to get the milli bugs as well as the ones with thin trunks so when the weather changed they died out. E: So what happened to the pilosas with thin trunks? P: They survived. E: What happened to the pilosas with wide trunks? P: They died out.
3	NS understanding in one generation	All of the following are present: - Sufficient knowledge of isolated facts - No misconception - Accurate mention of differential survival advantage - Accurate mention of differential reproduction in one	*Level 3 response: Differential survival and differential reproduction*[3] E …now most of the grown-up rudoos have longer necks. How do you think that happened? P: I don't know. E: What's your best guess? P: The ones with the shorter necks all died out because they couldn't reach the fruit and then the ones with the longer necks could reach the fruit and had more babies so there were more ones with longer necks.

		generation	E: What happened to the rudoos with longer necks? P: I don't know. E: What's your best guess? P: They could reach the fruit so they had more babies so there were more and more and more of them. E: Why? P: Because the fruit was up high and the little ones couldn't reach it, the ones with the short necks couldn't reach it, and the ones with the longer necks could reach the fruit.
4	NS understanding in multiple generations	All of the following is present: - Sufficient knowledge of isolated facts - No misconception - Accurate mention of differential survival advantage - Accurate mention of differential reproduction in one generation - Accurate mention of differential reproduction in multiple generations	*Level 4 response: Differential survival and reproduction in multiple generations*[2] E: …now okapis only have short necks. Why do you think that happened? P: The weather changed and the short neck okapis couldn't get any of the fruit that they need to live. E: What happened to the okapis with short necks? P: They probably died out. E: What happened to the okapis with long necks? P: They had babies and then these had babies and then they kept on having babies. *Level 4 response: Differential survival and reproduction in multiple generations*[3] E: …now most of the grown-up rudoos have longer necks. How do you think that happened? P: Um, you, these [points to shorter necks in past group] couldn't really eat a lot, and they died of starvation, and these [points to longer necks in past group] got a lot of, lot of things to eat, and had babies, and these [points to shorter necks in past group] mostly died out of starvation. E: What happened to rudoos with longer necks? P: Mmm, they live. E: And why do they live? P: Bec-c-… because they got enough food t-to eat. E: And so what happened next after they lived? P: …They had children and then died. E: And why is that? P: …because everything dies, and they ha-- they got children because they got a lot of, a lot of things to eat. E: And so what happened next after they had children and then died? P: Um, their children grew up to be grown-up rudoos, and then the same thing happened, like, they got old, they had children, and then they died. And the cycle…

Note. E = Experimenter; P = Participant. [1]Sufficient knowledge of isolated facts was defined as accurately answering and justifying 4 of 5 closed-ended questions in Experiment 1 and 5 of 6 closed-ended questions in Experiment 2. [2] Full open-ended responses taken from Experiment 1. [3]Open-ended responses taken from Experiment 2 (edited for length).