Journal of Experimental Child Psychology 203 (2021) 105041

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The effect of object similarity and alignment of examples on children's learning and transfer from picture books

Gabrielle A. Strouse [a,b,*], Patricia A. Ganea [c]

[a] Division of Counseling and Psychology in Education, Vermillion, SD 57069, USA
[b] Center for Brain and Behavior Research, University of South Dakota, Vermillion, SD 57069, USA
[c] Department of Applied Psychology and Human Development, Ontario Institute for Studies in Education, University of Toronto, Toronto, Ontario M5S 1V6, Canada

ARTICLE INFO

Keywords:
Analogical comparison
Object similarity
Structural alignment
Interleaving
Picture books
Transfer

ABSTRACT

Story picture books with examples can be used to teach young children science concepts. Learners can abstract relational information by comparing the analogical examples in the books, leading to a more abstract transferrable understanding of the concept. The purpose of this study was to determine whether manipulating the content or arrangement of the examples included in a picture book would support children's generalization and transfer of a relational concept, namely color camouflage. In total, 81 3-year-olds and 80 4-year-olds were read one of four books at two visits spaced approximately 1 week apart. Examples were manipulated in a 2 (Object Similarity: high or low) × 2 (Arrangement: interleaved or blocked) design. At each visit, children were asked forced-choice questions with photographs (generalization) and real animals (transfer) and needed to explain their choices. At the first visit, only 3-year-olds who had been read a book with high object similarity displayed generalization and transfer. After they were read the same book again at the second visit, 3-year-olds in all conditions performed above chance on generalization questions but made more correct selections if they had been read the books with blocked examples. The 4-year-olds showed no book-related differences on forced-choice questions at either visit but gave better explanations at the second visit if they had been read interleaved books. Our study provides evidence that picture books with

* Corresponding author at: Division of Counseling and Psychology in Education, Center for Brain and Behavior Research, University of South Dakota, Vermillion, SD 57069, USA.

E-mail address: gabrielle.strouse@usd.edu (G.A. Strouse).

https://doi.org/10.1016/j.jecp.2020.105041

analogical examples can be used to teach children about science but that different types and arrangements of examples may better support children at different ages and with different amounts of prior experience.

Introduction

Story picture books, designed for adults to share with prereaders, can be an effective means of teaching young children about science. Researchers have used picture books to teach young children both biological science information (Ganea, Canfield, Simons-Ghafari, & Chou, 2014; Ganea, Ma, & DeLoache, 2011; Gripshover & Markman, 2013; Kelemen, Emmons, Seston Schillaci, & Ganea, 2014; Tare, Chiong, Ganea, & DeLoache, 2010; Walker, Gopnik, & Ganea, 2015; Waxman, Herrmann, Woodring, & Medin, 2014) and physical science information (Ganea, Venkadasalam, & Larsen, in press; Ganea, Walker, & Simons-Ghafari, 2017; Venkadasalam & Ganea, 2018). Picture books may be effective for teaching young children scientific knowledge because they often include multiple concrete examples to illustrate scientific concepts (Ganea et al., 2011) and because providing visual depictions alongside verbal examples may help to reduce memory demands (Mayer & Moreno, 2003). Therefore, picture books are a promising avenue for teaching children scientific information prior to school entry.

However, learning scientific concepts from picture books can also be challenging. It may be difficult for children to identify the appropriate deep features of examples to incorporate into their concepts (see Strouse, Nyhout, & Ganea, 2018). It may also be difficult for children to transfer the newly acquired knowledge to contexts outside of the book. In some studies, children have been successful at transferring scientific knowledge from picture books to real-world contexts (Ganea et al., 2011; Ganea et al., in press; Strouse & Ganea, 2016; Venkadasalam & Ganea, 2018; Walker et al., 2015), but these studies have also shown limits to children's success. For example, Walker et al. (2015) found that 4- to 6-year-olds were more likely to generalize causal biological information to real situations when the story they learned from was realistic, and therefore more similar to the real-world situation, than when it was fantastical. In two other studies (Ganea et al., 2011; Strouse & Ganea, 2016), 4-year-olds transferred information about camouflage from a picture book to real animals, but performance was much higher on questions that mirrored the structure of the examples in the book than on questions that required children to think about the problem in a different way. In addition, some 4-year-olds could provide correct conceptual explanations for their answers, but performance was far from ceiling (Ganea et al., 2011; Strouse & Ganea, 2016), and explanations given by 3-year-olds were even more limited (Ganea et al., 2011). One remaining question is how to best structure the examples in the books to support children's conceptual understanding as well as promote transfer of their knowledge.

Existing literature has begun to address fantastical contexts and anthropomorphism in children's picture books (Ganea et al., 2014; Geerdts, 2016; Hopkins & Weisberg, 2017; Larsen, Lee, & Ganea, 2018; Strouse et al., 2018; Waxman et al., 2014) but has focused less on the content and structural arrangement of the specific examples provided. For the current study, we looked to the literature on analogical reasoning and learning through comparison to guide us in designing and testing changes in the content and arrangement of examples in picture books as a support for children's learning and transfer. We expanded on prior work in which children were taught the concept of biological color camouflage using picture books with examples of camouflaged and noncamouflaged animals (Ganea et al., 2011; Strouse & Ganea, 2016) by manipulating the content of the examples provided as well as their order of presentation within the book. This examination will lead to better understanding of how we can optimally design books to teach young children scientific concepts.

Comparing and contrasting examples

Relational categories are categories defined by relations between or among entities rather than intrinsic features of category members (Gentner, 2005). The concept of camouflage is relational because category membership depends on the relation between animals' attributes and their surrounding environment. Animals that have evolved specific camouflage mechanisms, such as having colors, shapes, or textures that frequently match the surroundings in which they live, are not actively engaged in camouflage at all times. Instead, the state of being camouflaged depends on whether their color, shape, or texture matches their current environment. Therefore, for children to identify whether an animal is actively color camouflaged, they need to learn to consider the relation between the animal's color and the color of its surrounding environment rather than the animal's color alone.

Learners acquire relational categories through a combination of language and experiences that allow them to identify the abstract relational structure of the category (Gentner, 2005). According to structure-mapping theory, these experiences often occur through the process of comparison, in which learners align the structure of examples by finding commonalities and differences among them (Gentner, 1983; Gentner & Markman, 1997; Markman, 1996; Markman & Gentner, 1996). Examples provided in picture books may support children's learning of relational categories by encouraging comparison with other previously encountered examples that learners have stored in memory. Through this act of comparison, learners identify relational commonalities between the examples, which supports them in abstracting relevant categorical structures (Gentner, 2005). For example, if children are shown an example of an arctic fox camouflaged in snow and they compare this example with their stored memory of a yellow butterfly on a yellow flower, they may abstract a general rule about the animal's color matching the color of its surroundings.

If learners have prior knowledge of the category, they may also compare the example with their existing knowledge of the rules or patterns associated with that category (Gentner, 2005). For example, if children have already abstracted the above rule about camouflage, they may use the arctic fox example to confirm its accuracy. However, if children's existing rule is associated with butterflies matching the color of flowers, the arctic fox example would not fit and children would be prompted to identify other relational abstractions to update their categorical structure.

Comparison of examples enables learners to refine their abstraction of relevant causal relations, which leads them to perceive these relations in other contexts and facilitates generalization and transfer to new situations (Goldwater & Gentner, 2015; Kurtz, Boukrina, & Gentner, 2013). Thus, children who have developed accurate abstract rules about camouflage through comparison will be able to identify camouflaged animals in new contexts.

Contrasting examples may also help children to develop or refine their abstract rules about a category's structure. Schwartz, Chase, Oppezzo, and Chin (2011) suggested that instructors may provide learners with contrasting cases that differ on one dimension to help them identify that dimension as an important part of the deep structure of the examples. Corral, Kurtz, and Jones (2018) showed that including contrasting examples of similar categories could help adults to learn a relational category, which they argued was because contrasts prompted the learners to identify diagnostic differences between the categories. In our case, providing examples of noncamouflaged animals that are highly similar to examples of camouflaged animals but differ in their color match with their background could help children to identify color match as a key feature or diagnostic difference that defines animals as camouflaged or not.

So far, contrasting examples have been used to teach children about category membership with mixed success. Namy and Clepper (2010) found that 4-year-olds who were visually presented with contrasting examples and left to draw their own contrasts did not appear to use the exemplars to rule out any criteria as a basis for category membership. Hammer, Diesendruck, Weinshall, and Hochstein (2009) found this also to be true of 6- to 9-year-olds; children in their study used similar visual examples to determine category membership but were unsuccessful at determining category membership when presented with contrasting examples. Yet in other research, preschoolers have successfully used information from contrasting examples. Kalish and Lawson (2007) found that 4- and 5-year-olds used both positive verbal evidence (e.g., "This one is black and is a raven") and negative verbal evidence (e.g., "This one is not black and is not a raven") when projecting biological attributes to new category exemplars. Similarly, Waxman, Lynch, Casey, and Baer (1997) found that 3- and 4-year-olds used verbally presented contrasting cases to help them establish subordinate categories and use them to make inferences. Thus, providing verbal scaffolds that highlight the key contrasting features of the exemplars, rather than leaving children to notice these features on their own, may help preschool children to use contrasts to develop and refine their knowledge about categories. The combination of verbal scaffolds alongside pictures is typical in educational picture books, and many of the lab studies using picture books to teach scientific knowledge to young children have paired exemplars with explanatory text to support children in identifying key features (Ganea et al., 2011; Kelemen et al., 2014; Strouse & Ganea, 2016; Venkadasalam & Ganea, 2018; Walker et al., 2015). In the current study, we followed the same pattern and created picture books that included both similar and contrasting cases paired with explanatory text.

Age differences in learning from examples

Children's efficiency in using picture book examples to identify relational commonalities and differences may differ across development. According to the relational shift hypothesis, young children often focus on object similarities, rather than relational similarities, when comparing examples (Ratterman & Gentner, 1998). As children acquire more relevant experience in a domain, they shift to a more relational focus (Kotovsky & Gentner, 1996; Pierce & Gholson, 1994; Ratterman & Gentner, 1998). That is, they look for abstract relational similarities, rather than surface-level object similarities, between examples. A similar shift can be seen in adults; those with greater domain knowledge are more likely than novices to categorize examples based on their structure than on their content (Goldwater & Gentner, 2015; Grégoire, Barr, & Shepherd, 2010). When children are young or inexperienced, their lack of attention to relational similarities may lead them to miss information important for generalization and transfer to new domains such as causal information (Goldwater & Gentner, 2015; Ratterman & Gentner, 1998). Learners with more domain knowledge, and therefore more attention to relational similarities, tend to be better at analogical transfer in that domain.

There is also evidence for developmental differences in how quickly children learn to solve analogical tasks. Ratterman and Gentner (1998) found that 4- and 5-year-olds fared well on a task that involved identifying relational similarities in size (i.e., clay pots decreasing in size), but many 3-year-olds needed direct instruction to scaffold their success on the task. Brown and Kane (1988) also found that 3-year-olds needed scaffolding on several tasks to prompt them to look for relational similarities between problems either through experiencing multiple prompts to transfer or by being asked to explain problem solutions. However, 4- and 5-year-olds needed just one example and no prompting prior to transferring the problem solution to a new context. These patterns were consistent even when children had no prior domain knowledge, which Brown and Kane (1988) interpreted to reflect developmental changes in children's use of learning strategies and metacognitive processes. Essentially, older children were more likely to look for and test abstract generalizable rules.

Differences in children's domain knowledge and their tendency to look for relational similarities and generalizable rules may help to explain why Ganea et al. (2011) found developmental differences in learning and transfer between 3- and 4-year-olds who were read a picture book about color camouflage. In the current study, we investigated two ways of presenting examples (object similarity and placement of contrasting examples) to support 3- and 4-year-olds in learning about camouflage, as demonstrated by their choices and explanations on generalization and transfer tasks. In the next sections, we review the literature related to these manipulations and explain why we expected to see developmental differences in how the presentation of examples would support children's performance on our tasks.

Object similarity

Examples that contain surface similarities support learners in making comparisons (Gelman, Raman, & Gentner, 2009; Gentner & Toupin, 1986). In the case that learners must draw examples or rules from memory to compare with a current example, surface similarities may act as cues that help learners to identify and retrieve appropriate information from memory (Gentner & Markman, 1997; Novick, 1988). Surface similarities help learners to align the examples structurally, compare them, and identify both their commonalities and differences (Gelman et al., 2009; Gentner & Gunn, 2001). Although experts may be more likely than novices to identify structural commonalities in examples that share few surface features (Novick, 1988), both adults and children appear to benefit from surface similarities when making comparisons in a wide variety of domains (Gentner, Lowenstein, & Hung, 2007; Gentner & Toupin, 1986; Haryu, Imai, & Okada, 2011; Holyoak & Koh, 1987).

We expected that greater surface similarities in picture book examples would also support children in learning about color camouflage regardless of age. For example, children learning from photos of similar lizards on similar backgrounds that differ only in color would be supported in aligning the examples and identifying the color match between the lizard and background as the key rule that defines color camouflage. However, children learning from photos of many different animals on different backgrounds would not as easily be able to identify the key relational similarities and differences in the photos. Therefore, we predicted that using books with examples that share surface features (i.e., similar looking lizards) would help children to identify a correct generalizable rule about camouflage that would further support them in later tests of generalization and transfer.

Arrangement of examples

Another way children may be supported in aligning examples is through copresentation of multiple examples. When examples are presented sequentially, learners must retrieve a relevant example or rule from memory to compare with new examples in order to identify relevant relational similarities. This could be especially challenging for young children who may have limited working memory and little relevant experience (Gentner, 2005). One way to address this problem is through joint presentation of multiple examples that invite learners to compare the examples with one another such that they are not required to select and retrieve a relevant example or rule from memory (Christie & Gentner, 2010). In picture books, examples can be presented side by side on opposing pages. Our goal was to examine whether same-category or contrasting examples were best presented side by side.

The arrangement of examples we used can be referred to as blocked (i.e., grouping same-category examples together) or interleaved (i.e., alternating examples from one category with examples from a different category). For adults, interleaving exemplars typically leads to better learning and discrimination between cases than blocking (Rohrer, 2012). According to study phase retrieval theory (Thios & D'Agostino, 1976), this is because interleaving requires learners to retrieve prior presentations from memory, which strengthens future retrievability and slows forgetting. That is, interleaving examples from different categories requires learners to search for a prior categorical rule or case for comparison each time an example is presented, and this search and retrieval process strengthens neural connections related to those comparisons. However, young children might not be able to retrieve the appropriate prior examples from memory, leading to little or no learning from interleaved examples (Vlach & Johnson, 2013). Blocking same-category examples and copresenting them should reduce memory demands, allowing children more opportunities to draw comparisons between category members. Therefore, blocked examples may support learning more than interleaved examples when prior categorical knowledge or working memory is limited.

The current study

We used a factorial design to experimentally study 3- and 4-year-olds' generalization and transfer of information about color camouflage acquired through examples in a picture book. We manipulated the object similarity of the examples and the arrangement (blocking or interleaving) of contrasting examples. We included multiple outcomes across multiple visits so that we could better understand the robustness of children's knowledge and identify change over time in both age groups. Because one goal of educational picture books is for children to acquire knowledge that they can transfer to real-world contexts, this included testing children's transfer to live animals (see also Ganea et al., 2011). We also included a pretest so that we could exclude children who already demonstrated knowledge of the relational category. We hypothesized that both 3- and 4-year-olds would benefit from examples with high object similarity. However, based on differences between 3- and 4-year-olds' analogical transfer in prior studies (Brown & Kane, 1988; Ganea et al., 2011; Ratterman & Gentner, 1998), we expected that different arrangements of examples may support younger and older children.

We hypothesized that 3-year-olds, who may need more support to look for similarities, have less prior experience with color camouflage, and have less available working memory to hold prior examples in mind than 4-year-olds, would benefit from blocked copresentation of same-category (noncamouflaged) examples, followed by blocked copresentation of camouflaged examples. The copresentation of same-category examples would serve to invite comparison and reduce the memory demands of doing so, providing a supportive context for 3-year-olds to develop an abstract relational rule to define the color camouflage category. In addition, because Brown and Kane (1988) found that 3-year-olds needed multiple requests to transfer or explain before they looked for relational similarities, the rereading of the same book on the second visit offered 3-year-olds the opportunity to deploy a new learning strategy at the second visit.

Alternatively, we hypothesized that 4-year-olds, who may have more prior knowledge, greater working memory capacity, and more inclination to use comparison as a strategy, would benefit from the interleaved examples, especially when contrasting objects were highly similar. The copresentation of different-category examples would invite contrasts, and high object similarity would help children to isolate color match as a defining feature of color camouflage.

Method

Participants

Participants were 81 3-year-olds (M = 41.37 months, SD = 3.54; 43 girls) and 80 4-year-olds (M = 53.38 months, SD = 3.61; 37 girls) from a large metropolitan area in North America (Table 1). Participants were recruited from a database of families who had expressed interest in research at local street fairs or responded to advertisements. An additional 15 children were not included in the analysis due to unwillingness to complete the study (n = 6), giving camouflage explanations at pretest (n = 4), low English proficiency (n = 3), or experimenter error (n = 2).

Parents identified their children's ethnicity as White (63%, n = 102), African Canadian (1%, n = 1), White Hispanic (1%, n = 2), Asian (6%, n = 9), Native American (1%, n = 1), other (5%, n = 8), or multiple (18%, n = 29) (9 parents declined to provide ethnicity information). Approximately half of the children in the study were from monolingual English-speaking homes. Of those who reported multiple languages spoken in the home, 90% reported English as the main language spoken. For those who reported a main language other than English, children's English proficiency was confirmed via parent report and performance on the Peabody Picture Vocabulary Test (PPVT; standard scores ranged from 96 to 126). Parents' median and modal education level was a 4-year college degree.

Materials

Picture books

Children were read one of four picture books (Open Science Framework: https://osf.io/ydm6s) by a researcher in a 2 (Object Similarity: high or low) × 2 (Arrangement: interleaved or blocked) design. Each book contained a narrative about a "bird" (hawk) that was hungry and looking for something

Table 1
Participant information with mean age in months by age group and condition.

	3-year-olds		4-year-olds	
Condition	n	Age (months)	n	Age (months)
High object similarity, interleaved	20	M = 41.90, SD = 3.37	20	M = 54.00, SD = 4.10
Low object similarity, interleaved	21	M = 40.60, SD = 3.43	20	M = 53.11, SD = 3.10
High object similarity, blocked	20	M = 41.60, SD = 3.63	20	M = 53.61, SD = 3.61
Low object similarity, blocked	20	M = 41.40, SD = 3.87	20	M = 52.79, SD = 3.71

to eat. Each page contained a photograph of a real animal in a natural environment. Over the course of the story, the hawk flew past eight different animals, four of which were color camouflaged in their environment and four of which were colored differently from their environment. The book text explained that the bird "could not see" each color camouflaged animal and kept flying but that it was "easy to see" each noncamouflaged animal and the bird would catch it. For consistently with prior studies, and to facilitate children's use of relational language rather than new vocabulary in their explanations, the word "camouflage" was not used. All books contained similar text, but the order of the pages and the pictures displayed alongside the text were manipulated (Fig. 1). The books were manipulated across two dimensions: similarity of the target animals displayed (high or low object similarity) and arrangement of the examples (blocked or interleaved).

The two books with high *object similarity* contained eight pictures of lizards; the two books with low object similarity contained two pictures of lizards, two squirrels, two frogs, and two butterflies. In the books with low object similarity, each type of animal appeared in one noncamouflaged instance and one camouflaged instance. The object similarity manipulation involved changing the pictures provided in the text—not the order of the pages or the text itself—although the animal names and

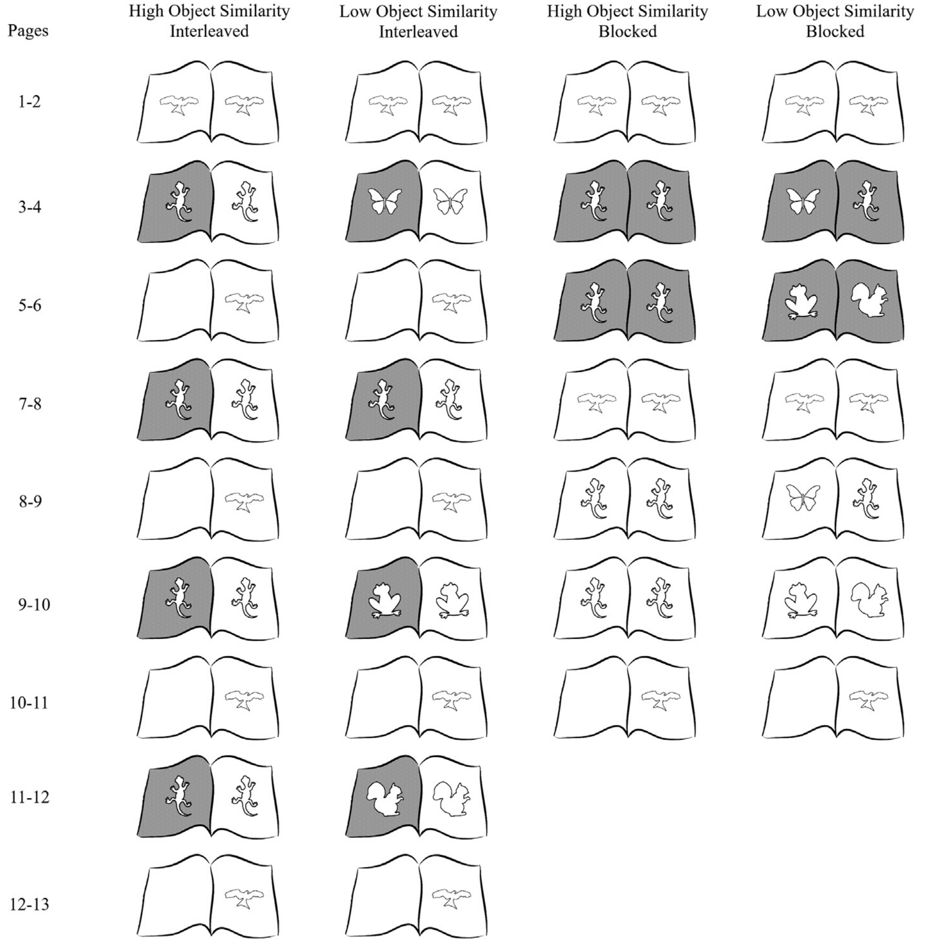

Fig. 1. Picture book layout in each condition. Shaded background represents noncamouflaged pages, and white background represents camouflaged pages. A copy of the images used in the study can be accessed at the Open Science Framework (https://osf.io/ydm6s).

locations in the text were modified when appropriate to match the pictures (e.g., "A green *lizard/frog* is on (*a*) green *leaf/grapes*").

The two *interleaved* books contained the same language and pictures as their counterparts in the *blocked* conditions (although one page of the hawk flying was removed). In the interleaved books, each left-hand page featuring a noncamouflaged animal was paired with a right-hand page featuring a camouflaged animal. Thus, children in the interleaved conditions could view both a noncamouflaged animal photograph and a camouflaged animal photograph at the same time. In the blocked books, the four noncamouflaged animals appeared first (in two side-by-side pairs) and the four camouflaged animals followed. Thus, the interleaving/blocking manipulation involved changing the order of the pages (see Fig. 1)—not the pictures or text.

Test items

Children were tested using a photograph of a hawk that was placed between two test items from which children were asked to select. At pretest, these items were photographs of a camouflaged caterpillar and a noncamouflaged caterpillar for the first trial and photographs of a camouflaged rabbit and a noncamouflaged rabbit for the second trial. At posttest, the hawk was placed between photographs of a camouflaged lizard and a noncamouflaged lizard that did not appear in the books for the first generalization trial and between photographs of a camouflaged crab and a noncamouflaged crab for the second trial (all photographs at https://osf.io/ydm6s). Children were then asked to respond to four additional questions using real animals in aquarium tanks; a pair of live leopard geckos and a pair of Corydoras catfish were used for these questions.

Procedure

Children were tested in a child-oriented research space on campus.

Visit 1

Pretest phase. The researcher, sitting across a child-sized table from the child, began by placing the hawk photograph in front of the child and stating, "This is a bird that eats caterpillars. It's very hungry and is looking for something to eat." She then placed the pair of caterpillar pictures (one camouflaged and one not) below the picture of the hawk and asked "Which caterpillar will the bird eat: this one or this one [pointing right to left]?" Once the child responded, the researcher said "Thank you" and asked the child to explain his or her selection. This procedure was then repeated with photographs of rabbits. The camouflaged animal appeared to the child's right on the first trial and to the child's left on the second trial. The 4 children who made two correct selections and gave camouflaged-based reasoning (i.e., received scores of 2) at pretest were excluded from further participation in the study.

Reading phase. The researcher then moved her chair so that she was seated next to the child at the table. She presented the book and said, "Now we're going to read a story." She proceeded to read the book verbatim without spontaneous remarks. If the child made comments or asked questions, she responded positively and briefly and then continued reading. The reading phase took approximately 3 min.

Posttest phase. Generalization. The posttest began with two photograph trials identical to those in the pretest with the exception of the photographs used. The goal of this test was to assess children's ability to generalize to novel exemplars—first to new exemplars of the same kind of animal that appeared in the books and then to exemplars of an animal that did not appear in the books. Children selected from photographs of two lizards on the first trial and two crabs on the second trial. Once children selected a picture, they were asked to explain their choice.

Transfer. This test involved four trials using live animals and assessed children's ability to transfer their knowledge from the picture book to a real-world context, that is, to see whether children could transfer knowledge across modality from a two-dimensional picture book to three-dimensional, live animals. In two *near transfer* trials, children were taken into an adjacent room and the hawk photograph was placed between two aquarium tanks: one containing a lizard the same color as the substrate and one containing a lizard a different color from the substrate. The camouflaged and noncamouflaged lizards were similar in size and length, and neither one was pictured in the books. Children were asked, "The bird is still hungry. It's looking for a lizard to eat. Which lizard will the bird eat: this one or this one [pointing to each lizard, right to left]?" Once children selected a tank, they were asked to explain their choice. Children were then taken to two aquariums with animals that did not appear in any of the books—fish—and the procedure was repeated. One tank contained a handful of same-color Corydoras catfish that matched the color of the rocks on the bottom, and the other tank contained Corydoras catfish that did not match the rocks. As the researcher asked the question and pointed to each tank (right to left), she pointed to a fish that was swimming close to the bottom of the tank. The camouflaged lizard appeared on children's right, and the camouflaged fish appeared on their left. These trials are referred to as near transfer because the wording and structure of this task were very similar to the wording and structure used in the book, where children could see animals in camouflaged and noncamouflaged contexts.

In two *far transfer* trials, children were taken into a different adjacent room and the hawk was placed between two aquarium tanks with different colored rocks. One of the lizards that had been used in the near transfer test was placed in a small carrying tank next to the hawk. Children were asked, "This bird is still hungry. It's looking for a lizard to eat. Where should we put the lizard so it can be safe from the bird: here or here [pointing right to left]?" Once children selected a tank, they were asked to explain their choice. Children were then taken to two different aquariums with rocks and water in them, and the procedure was repeated. Similarly, one of the fish from the near transfer test was placed in a small carrying tank, and children were asked which of the two aquariums to place it in so it would be safe. The rocks that matched the color of the animal appeared on children's right for the lizard and on their left for the fish. These questions are referred to as far transfer because the wording of this task was different from that used in the book, children were asked to select a background rather than an animal, and a correct answer involved an animal/background color match rather than a mismatch. In addition, rather than just needing to choose between two contexts that included the animals (as in the book), children needed to select a context in which to place one animal in order for it to be camouflaged. The order of the near and far transfer trial sets was counterbalanced, but lizard questions were always asked first.

Visit 2

A total of 149 children (76 3-year-olds and 73 4-year-olds) returned after a delay (*M* = 6.05 days, *SD* = 3.10, range = 1–18).

Reading. Children were read the same book in the same manner as on the first visit.

Posttest. Children completed the same generalization, near transfer, and far transfer trials as on Visit 1. Whether near or far transfer trials occurred first was counterbalanced.

Children also completed the PPVT (Dunn & Dunn, 2007), a test of receptive vocabulary. Children were asked to point to the picture (from a set of four pictures) that went with the prompt given by the researcher. Children's standard scores ranged from 84 to 145 (*M* = 117.8, *SD* = 12.6).

Coding

Children's choices on the pretest and posttest questions were coded as correct or incorrect by the experimenter at the time of participation. The responses were also coded from video by a second coder who was blind to condition. Approximately 88% of the responses were codable from the videotapes. Reliability was kappa = .87. Discrepancies were resolved by review of the videos by a third coder who was also blind to condition.

Children's explanations for their selections on the pretest and posttest questions were transcribed from video. Two coders blind to condition coded children's explanations for forced-choice trials on which they selected the correct animal or tank using the method described in Strouse and Ganea (2016). Children received a score of 0 if they did not mention color or visibility, 1 if they provided an incomplete response about color or visibility (e.g., referenced the color of only the lizard or the background), or 2 if they provided a complete response about camouflage (e.g., mentioned the matching color of both the lizard and the background). Reliability for explanations given across all posttest trials (generalization and transfer), assessed using a weighted kappa, was kappa = .88. Children were given a total reasoning score from 0 to 2 for each visit by averaging their responses across the generalization and transfer trials.

Results

Preliminary analyses indicated that the number of days between visits was not associated with children's forced-choice or reasoning scores at the second visit, so this variable was not included in further analyses.

Forced-choice questions

To examine the effect of book type on performance, children's responses to the forced-choice questions were analyzed using a logistic generalized estimating equation (GEE) model to account for the binary and repeated nature of the data. The model included a question set (generalization, near transfer, or far transfer) and trial within each set (lizards or crabs/fish) as repeated measures. In addition, object similarity, arrangement of examples (interleaved or blocked), Object Similarity × Arrangement, Object Similarity × Question Set, and Arrangement × Question Set were included as independent variables. None of the interaction terms led to significant results, and removing them increased model fit, so they were removed from the model. We also controlled for pretest performance and vocabulary in each model. We used an autoregressive covariance structure because we expected children's responses to be less correlated as the questions became less similar to the presentation format of the information in the book.

Visit 1

Analyses indicated that 3-year-old children made more correct choices if they had been read books with high object similarity, $\chi^2 = 4.39$, $p = .036$ (high object similarity: $M = .60$, $SE = .032$; low object similarity: $M = .52$, $SE = .032$). There was also a main effect of question set, $\chi^2(2) = 22.68$, $p < .001$. Post hoc comparisons indicated that performance on generalization and near transfer trials was significantly higher than performance on far transfer trials, *p*s < .001 (generalization: $M = .69$, $SE = .037$; near transfer: $M = .60$, $SE = .039$; far transfer: $M = .40$, $SE = .039$). There were no main effects of example arrangement, trial, pretest performance, or vocabulary.

The 4-year-old children also displayed a main effect of question set, $\chi^2(2) = 13.87$, $p = .001$. Post hoc analyses indicated that generalization performance and near transfer performance were significantly higher than far transfer performance, *p*s < .001 (generalization: $M = .89$, $SE = .025$; near transfer: $M = .78$, $SE = .033$; far transfer: $M = .70$, $SE = .037$). There were no main effects of object similarity, example arrangement, trial, pretest performance, or vocabulary.

Because there were no main effects of trial and no interactions between either condition manipulation and trial, children's performance across the two trials in each question set were collapsed for analyses against chance. Results are presented in Fig. 2. In general, the 4-year-olds scored above chance on generalization and near transfer trials regardless of the book they were read. The 3-year-olds scored above chance rarely and only if they were read books with high object similarity.

Visit 2

At the second visit, 3-year-old children made more correct choices if they were read books with blocked examples, $\chi^2(1) = 6.46$, $p = .011$ (interleaved: $M = .50$, $SE = .034$; blocked: $M = .62$, $SE = .031$). There was a main effect of question set, $\chi^2(2) = 45.42$, $p < .001$. Post hoc tests indicated that children's scores on all trial sets significantly differed from one another (generalization: $M = .77$, $SE = .034$; near transfer: $M = .53$, $SE = .041$; far transfer: $M = .39$, $SE = .040$). There were no main effects of object similarity, trial, or vocabulary. Pretest performance was negatively associated with posttest performance, $\chi^2(2) = 6.47$, $p = .039$.

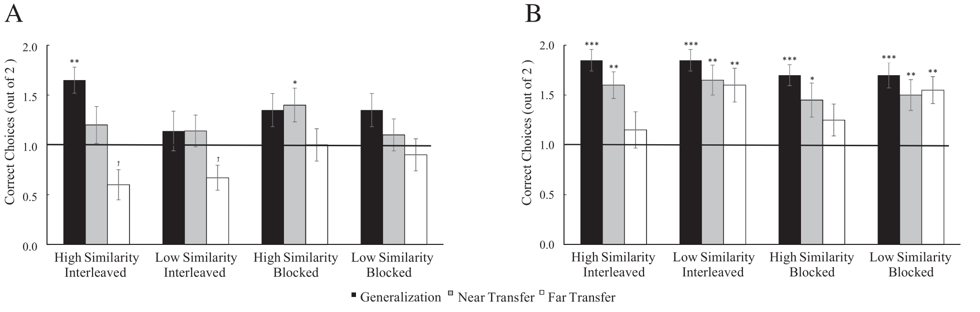

Fig. 2. Correct responses by 3-year-olds (A) and 4-year-olds (B) to forced-choice questions on Visit 1. Means and standard errors by condition are shown. Asterisks indicate that scores (out of 2) were significantly above a chance value of 1 as indicated by a Wilcoxon signed-rank test: ***$p < .001$; **$p < .01$; *$p < .05$. †Value is significantly below chance at $p < .05$.

The 4-year-olds again displayed a main effect of question set, $\chi^2(2) = 7.81$, $p = .020$. Post hoc analyses indicated that generalization performance was significantly higher than both near transfer and far transfer performance, *ps* < .025 (generalization: $M = .82$, $SE = .032$; near transfer: $M = .72$, $SE = .037$; far transfer: $M = .70$, $SE = .038$). There were no main effects of object similarity, arrangement of examples, trial, pretest performance, or vocabulary.

Tests against chance are presented in Fig. 3. In general, both 3- and 4-year-olds performed above chance on generalization trials. The 3-year-olds did not show evidence of transfer in any condition, and the 4-year-olds did so only sporadically.

Camouflage-based reasoning

Children's explanations for their choices (coded on a scale of 0–2) were analyzed using factorial analyses of covariance (ANCOVAs) with the condition manipulations (object similarity and arrangement of examples) and age group as the independent variables and children's standardized PPVT scores as a continuous covariate. Bootstrapping with 1000 samples was used to address any concerns about violations of distributional assumptions.

Visit 1

There were no significant differences in children's explanations by condition at Visit 1 and no significant interactions. The 4-year-olds, $F(1, 137) = 59.68$, $p < .001$, $\eta^2 = .28$, and children with higher

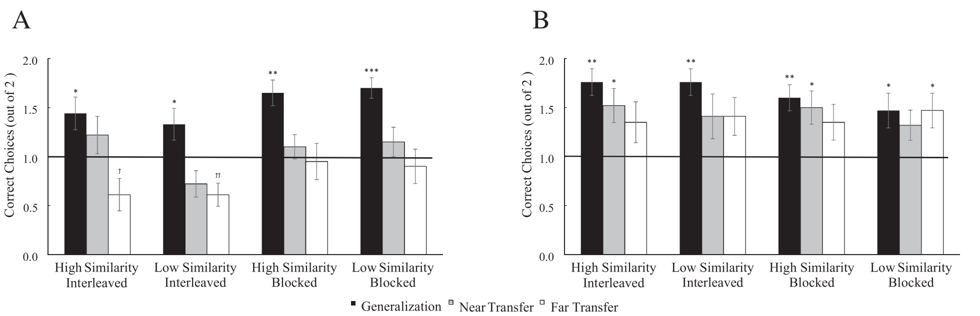

Fig. 3. Correct responses by 3-year-olds (A) and 4-year-olds (B) to forced-choice questions on Visit 2. Means and standard errors by condition are shown. Asterisks indicate that scores were significantly above a chance value of 1 as indicated by a Wilcoxon signed-rank test: ***$p < .001$; **$p < .01$; *$p < .05$. †Value is significantly below chance at $p < .05$. ††Value is significantly below chance at $p < .01$.

vocabulary scores, $F(1, 137) = 9.66$, $p = .002$, $\eta^2 = .05$, gave higher-scoring explanations. Mean scores are presented in Fig. 4.

Visit 2

At Visit 2, 4-year-olds, $F(1, 132) = 75.32$, $p < .001$, $\eta^2 = .32$, and children with higher vocabulary scores, $F(1, 132) = 7.17$, $p = .008$, $\eta^2 = .03$, again gave higher-scoring explanations. There was also a main effect of the arrangement of examples, $F(1, 132) = 7.34$, $p = .008$, $\eta^2 = .03$, and an Age × Arrangement interaction, $F(1, 132) = 5.73$, $p = .018$, $\eta^2 = .03$. Simple-effects analysis revealed that the effect of arrangement was significant only for 4-year-olds, $F(1, 132) = 12.86$, $p < .001$ [3-year-olds: $F(1, 132) = 0.10$, $p = .759$]. The 4-year-olds provided better explanations at the second visit if they were read the interleaved books rather than the blocked books. There were no other significant main effects or interactions.

Discussion

We investigated the effects of two manipulations of the examples provided in a picture book on 3- and 4-year-olds' generalization and transfer of knowledge about the relational category of color camouflage. After having been read the book once on the first visit, the 3-year-olds displayed generalization and transfer only if they had been read the books with high object similarity. The 4-year-olds displayed generalization and transfer of categorical knowledge regardless of which book they were read. There were no condition-related differences in the quality of explanations provided by children in either age group.

After they were read the same book again on a second visit, the 3-year-olds in all conditions performed above chance on generalization questions about color camouflage, with those who were read books with blocked examples outperforming the others by answering significantly more questions correctly. The 4-year-olds again displayed no significant condition differences in generalization and transfer of categorical knowledge. However, analysis of their explanations for test choices revealed that 4-year-olds provided more complete color camouflage explanations if they had been read the interleaved books.

Object similarity

We hypothesized that both 3- and 4-year-olds would benefit from examples with high object similarity. Consistent with our hypothesis, 3-year-olds scored higher on the generalization and transfer tests at the first visit if they had been read the books with high object similarity rather than the books with low object similarity. This is consistent with benefits from object similarity found in other studies (Gentner et al., 2007; Gentner & Toupin, 1986; Haryu et al., 2011) and is consistent with the above-chance generalization performance of 3-year-olds in a prior study where books with high object

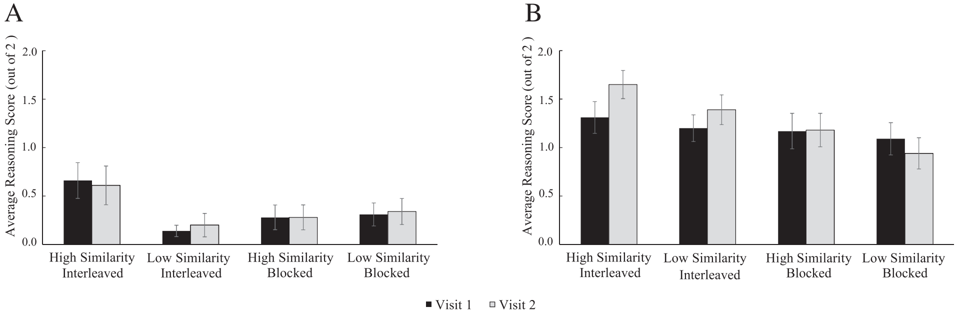

Fig. 4. Reasoning scores of 3-year-olds (A) and 4-year-olds (B) at each visit by condition. Children were given a total reasoning score out of 2 by averaging the scores given to their explanations across all posttest trials.

similarity were used (Ganea et al., 2011). However, contrary to our hypothesis, there were no benefits to having been read a book with high object similarity for 3-year-olds at their second visit or for 4-year-olds on either visit. In these cases, 3- and 4-year-olds showed consistent above-chance generalization from all books, and 4-year-olds showed some transfer to real animals regardless of the book they were read. One interpretation is that the books with high object similarity helped the 3-year-olds develop initial generalizable rules about camouflage at the first visit, but that other factors were more important for supporting children once they had more experience. In prior studies, 4-year-olds were more likely than 3-year-olds to look for and transfer relational similarities on their own without prompting (Ganea et al., 2011; Ratterman & Gentner, 1998), and 3-year-olds began to do so after being prompted with transfer questions (Brown & Kane, 1988). It is possible that the books with high object similarity supported 3-year-olds' performance at the first visit by prompting children to begin the process of searching for relational similarities prior to being asked the first set of transfer questions. The 3-year-olds may have also benefited from answering a question about lizards directly after having seen the lizard examples in the book.

We also hypothesized that object similarity might be particularly important for children in the interleaved condition to increase the alignment across the contrasting examples and better isolate the animals' color match with its surroundings as a diagnostic difference. This would be consistent with Rohrer's (2012) prediction, that when examples were less similar, the benefit of interleaving would be smaller. Although children displayed the most camouflage-based reasoning in their explanations if they had read the interleaved books with high object similarity, this interaction was nonsignificant. It is possible that the abstract rules that children developed from highly similar examples were not refined enough to support them in articulating these rules or applying them to new situations. It is also possible that the interaction between object similarity and interleaving would have been more substantial had we used different animals both within and across the pairs of examples in the interleaved books with low object similarity.

According to Gentner and Hoyos (2017), breadth of training promotes breadth of transfer for relational learning only when differences are alignable. That is, when a learner *successfully* compares very different examples, this leads to abstract generalizable rules that promote transfer. However, their success may depend on the examples having enough similarity to support alignment. In the case of our study, books with high object similarity may have led to more successful comparisons but also a more context-bound understanding of camouflage than books with low object similarity. These two competing mechanics could have essentially cancelled each other out, resulting in a lack of condition differences on many of our tests. A better support for learning may have been to provide high object similarity first to support children's initial relational comparisons, followed by low object similarity to support more generalized abstractions. Therefore, future studies might test whether children would benefit from being read books with high object similarity first, followed by books with more diverse examples, especially if variations in the examples are similar to the type of variations children will encounter on later transfer tasks.

Arrangement of examples

We hypothesized that 3-year-olds, who may need more support to prompt them to look for similarities and have less prior experience with color camouflage and less available working memory to hold prior examples in mind than 4-year-olds, would benefit from blocked copresentation of same-category (noncamouflaged) examples, followed by blocked copresentation of camouflaged examples. The copresentation of same-category examples would serve to invite comparison and reduce the memory demands of doing so, providing a supportive context for 3-year-olds to develop an abstract relational rule to define the color camouflage category. Our hypothesis was supported only at the second visit after children had experienced the transfer tasks at the first visit. This indicates that blocked copresentation alone might not have been enough to facilitate 3-year-olds in noticing relational similarities, but perhaps once they were prompted to look for similarities by the transfer tasks at the first visit, blocking supported them in doing so.

Alternatively, we hypothesized that 4-year-olds, who may have more prior knowledge, greater working memory capacity, and more inclination to use comparison as a strategy in the absence of explicit scaffolding, would benefit from the interleaved examples. We expected the interleaved condition to be the most difficult because placing contrasting examples side by side would facilitate learning through contrast (Rohrer, 2012) and because to compare same-category examples children would need to recall an example or rule from a prior page. This arrangement could have resulted in 4-year-olds in the interleaved conditions using both comparison *and* contrast to develop and refine their rules, potentially leading to the most refined abstract rules about color camouflage. We found support for this hypothesis in 4-year-olds' use of camouflage-based reasoning in the explanations they gave on the second visit. This is consistent with prior studies in which 3- to 5-year-olds who were presented with contrasts scaffolded by verbal labels were able to use them to make inferences about learned categories (Kalish & Lawson, 2007; Waxman et al., 1997).

Multiple visits

We expanded on prior studies by testing children across multiple visits to understand how their knowledge changed after multiple exposures to the books. In Ganea et al. (2011) study, 3-year-olds had some success at forced-choice generalization questions after being read their book about camouflage once, which included interleaved examples with high object similarity (camouflaged and noncamouflaged frogs). However, there was little evidence that 3-year-olds could use camouflaged-based reasoning to explain their responses. In the current study, 3-year-olds also had some success at the generalization questions at both visits, benefiting from books with high object similarity at the first visit and blocking at the second visit. Neither manipulation appeared to provide much support for their verbal reasoning or transfer to real animals. It is possible that 3-year-olds were beginning to notice relational patterns in the books' examples but could not yet articulate or transfer them.

In both similar prior studies (Ganea et al., 2011; Strouse & Ganea, 2016), 4-year-olds displayed generalization, transfer to real animals, and camouflage-based reasoning after being read books with high object similarity and interleaving one time. Our 4-year-olds also generalized, transferred, and used camouflage-based reasoning at both visits regardless of which book they were read. In addition, 4-year-olds displayed even greater levels of camouflage-based reasoning at the second visit if they had been read the interleaved books, suggesting that interleaving across multiple readings supported them in articulating their abstract rules.

Limitations and future directions

In this study, the word "camouflage" was never used to describe the animals depicted in the books. This choice was made to be consistent with prior research (Ganea et al., 2011; Strouse & Ganea, 2016) and because we were interested in the reasons children would give for their choices, rather than the learning of the correct label for the concept. However, linguistic labels have been shown to invite comparison and promote relational abstraction (Christie & Gentner, 2014; Gentner & Christie, 2010; Gentner & Namy, 1999). Therefore, the inclusion of the word "camouflage" to label in-category examples and distinguish noncamouflage examples as out of category may support children in more efficiently abstracting rules about the category. Future researchers should consider testing the effect of labels in conjunction with variations in the way examples are presented on children's learning of the conceptual category of camouflage.

Children in this study were read the same book and completed the same testing procedures at both visits. Testing was included at the first visit to allow for direct comparisons with prior studies and provide information about children's knowledge at both time points. However, because children participated in both reading and testing at the first visit, we cannot separate the effects of these activities on children's performance at the second visit. Although we did not directly test differences in performance from the first visit to the second visit, tests against chance indicated some improvement in 3-year-olds' generalization and some decline in 4-year-olds' transfer. Increases may have been a result of children's increased exposure to the books, consolidation of learning during the delay, or other factors. Participation in the testing procedure at the first visit may have led to both increases in performance (e.g., by cueing children to pay attention to relevant details) and decreases (e.g., far transfer questions disrupting a learned rule to always choose the color mismatch). Future researchers may wish to separate these factors by eliminating testing on the first visit or including a condition in which children are not re-exposed to the books on the second visit.

We made two choices about the order in which examples were presented that may have been important to our pattern of results. First, we chose to present the noncamouflaged examples first, for consistency across conditions and to decrease the complexity of our design, which already included many counterbalanced factors. It is possible that seeing the examples of animal/background color mismatch prior to the examples of color match influenced children's identification and reasoning about the relational category. This could be empirically tested in future studies. Second, we chose to block or interleave individual examples rather than pairs of examples. As a result, blocking facilitated comparison through joint presentation of same-category examples, whereas interleaving facilitated contrast through joint presentation of different-category examples. Alternatively, future researchers could choose to block or interleave individual examples sequentially (with a page flip in between) or interleave pairs of examples such that pairs of same-category examples would alternate with one another (e.g., a noncamouflage–noncamouflage pair followed by a camouflage–camouflage pair). This would help to clarify the importance of comparison and contrast as a mechanism for explaining differences in learning from blocked and interleaved examples.

Finally, we did not measure children's inhibitory control or working memory. Several studies have suggested that there may be differences in analogical reasoning performance in cases where working memory or inhibitory control is limited (Morrison et al., 2004; Richland, Morrison, & Holyoak, 2006; Simms, Frausel, & Richland, 2018; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004; Waltz, Lau, Grewal, & Holyoak, 2000). One explanation may be that cognitive resources are needed to suppress irrelevant matches (Morrison, Doumas, & Richland, 2011). In our study, failure to suppress irrelevant responses may partially explain children's consistently lower (and sometimes significantly below-chance) performance on the far transfer questions. The far transfer questions were included in this study as an opportunity for children to apply camouflage-based reasoning in a new context. However, to successfully answer these questions, children needed to suppress choosing the animal/background mismatch, which had been the correct answer to prior questions and was consistent with the way the examples were presented in the book. Children instead needed to choose the matched animal/background in response to the question about which animal would be safe. The inclusion of inhibitory control or working memory measures in future studies may help to tease apart whether children's low performance on far transfer questions was due to their limited understanding of camouflage or an inability to suppress irrelevant responses.

Summary

Object similarity supported initial development of the relational category of color camouflage for the 3-year-olds, but this advantage did not extend to continued higher performance after the second reading and testing session. Instead, blocking of same-category examples facilitated 3-year-olds' generalization on the second visit after they had prior exposure to the book and had attempted to solve the transfer questions. These children had limited ability to transfer or explain their choices at either visit. The 4-year-olds generalized, transferred, and used camouflage-based reasoning at both visits regardless of which book they were read. Interleaved examples led to higher reasoning scores after two reading and testing sessions, possibly because this group was able to benefit from both comparing and contrasting examples to develop and test their abstract rules.

Conclusion

Scientists use analogy to scaffold the development of new knowledge and help them explain scientific concepts to others (Dunbar, 1997). Parents and teachers also naturally use analogy when they teach children about science (Richland, Holyoak, & Stigler, 2004; Tare, French, Frazier, Diamond, & Evans, 2011; Valle & Callanan, 2006). Vendetti, Matlen, Richland, and Bunge (2015) argued that analogy is a good way for educators to teach children but that young children may need more explicit support for learning from analogy than older children and adults. Our study provides evidence that picture books with analogical examples can be used to teach children about science but that different types and arrangements of examples may better support children at different ages and with different amounts of prior experience.

CRediT authorship contribution statement

Gabrielle A. Strouse: Investigation, Formal analysis, Writing - original draft, Project administration. **Patricia A. Ganea:** Conceptualization, Methodology, Writing - review & editing, Supervision, Funding acquisition.

Acknowledgments

Thank you to the families who participated in the research and to Heather Gallant for data collection, Sharon Chan and Savannah Barker for data coding, Susan Strouse for illustrations used in Fig. 1, and Laura Weinheimer and others in the Language and Learning Lab for helpful feedback during design of the study. This work was supported by an Insight Development Grant from the Social Sciences and Humanities Research Council of Canada.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2020.105041.

References

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.
Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*, 356–373.
Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*, 383–397.
Corral, D., Kurtz, K. J., & Jones, M. (2018). Learning relational concepts from within- versus between-category comparisons. *Journal of Experimental Psychology: General, 147*, 1571–1596.
Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461–493). Washington, DC: American Psychological Association Press.
Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Minneapolis, MN: Pearson Assessments.
Ganea, P. A., Canfield, C. F., Simons-Ghafari, K., & Chou, T. (2014). Do cavies talk? The effect of anthropomorphic picture books on children's knowledge about animals. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00283.
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.
Ganea, P. A., Larsen, N. E., & Venkadasalam, V. P. (in press). The role of alternative theories and anomalous evidence in children's scientific belief revision. Child Development.
Ganea, P. A., Walker, C. M., & Simons-Ghafari, K. (2017, April). Weighing the evidence promoting belief revision through storybooks. Paper presented at the biennial meeting of the Society for Research in Child Development. Austin, TX.
Geerdts, M. S. (2016). (Un)real animals: Anthropomorphism and early learning about animals. *Child Development Perspectives, 10*, 10–14.
Gelman, S. A., Raman, L., & Gentner, D. (2009). Effects of language and similarity on comparison processing. *Language Learning and Development, 5*, 147–171.
Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.
Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 263–294). New York: Psychology Press.
Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition, 2*, 261–283.
Gentner, D., & Gunn, V. (2001). Structural alignment facilitates the noticing of differences. *Memory & Cognition, 29*, 565–577.
Gentner, D., & Hoyos, C. (2017). Analogy and abstraction. *Topics in Cognitive Science, 9*, 672–693.
Gentner, D., Lowenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*, 285–307.
Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56.
Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513.
Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.
Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition, 137*, 137–153.

Grégoire, D. A., Barr, P. S., & Shepherd, D. A. (2010). Cognitive processes of opportunity recognition: The role of structural alignment. *Organization Science, 21*, 413–431.
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.
Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference?. *Cognition, 112*, 105–119.
Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development, 82*, 674–686.
Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332–340.
Hopkins, E. J., & Weisberg, D. S. (2017). The youngest readers' dilemma: A review of children's learning from fictional sources. *Developmental Review, 43*, 48–70.
Kalish, C. W., & Lawson, C. A. (2007). Negative evidence and inductive generalisation. *Thinking & Reasoning, 13*, 394–425.
Kelemen, D., Emmons, N. A., Seston Schillaci, R., & Ganea, P. A. (2014). Young children can be taught basic natural selection using a picture-storybook intervention. *Psychological Science, 25*, 893–902.
Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.
Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1303–1310.
Larsen, N. E., Lee, K., & Ganea, P. A. (2018). Do storybooks with anthropomorphized animal characters promote prosocial behaviors in young children?. *Developmental Science, 21*, e12590.
Markman, A. B. (1996). Structural alignment in similarity and difference judgments. *Psychonomic Bulletin & Review, 3*, 227–230.
Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition, 24*, 235–249.
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52.
Morrison, R. G., Doumas, L. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science, 14*, 516–529.
Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., & Knowlton, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience, 16*, 260–271.
Namy, L. L., & Clepper, L. E. (2010). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology, 107*, 291–305.
Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 510–520.
Pierce, K. A., & Gholson, B. (1994). Surface similarity and relational similarity in the development of analogical problem solving: Isomorphic and nonisomorphic transfer. *Developmental Psychology, 30*, 724–737.
Ratterman, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal mapping task. *Cognitive Development, 13*, 453–478.
Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). Analogy use in eighth-grade mathematics classrooms. *Cognition and Instruction, 22*, 37–60.
Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249–273.
Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educational Psychology Review, 24*, 355–367.
Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*, 759–775.
Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology, 166*, 160–177.
Strouse, G. A., & Ganea, P. A. (2016). Are prompts provided by electronic books as effective for teaching preschoolers a biological concept as those provided by adults?. *Early Education and Development, 27*, 1190–1204.
Strouse, G. A., Nyhout, A., & Ganea, P. A. (2018). The role of book features in young children's transfer of information from picture books to real-world contexts. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00050.
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*, 395–400.
Tare, M., French, J., Frazier, B. N., Diamond, J., & Evans, E. M. (2011). Explanatory parent–child conversation predominates at an evolution exhibit. *Science Education, 95*, 720–744.
Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15*, 529–536.
Valle, A., & Callanan, M. A. (2006). Similarity comparisons and relational analogies in parent–child conversations about science topics. *Merrill-Palmer Quarterly, 52*, 96–124.
Vendetti, M. S., Matlen, B. J., Richland, L. E., & Bunge, S. A. (2015). Analogical reasoning in the classroom: Insights from cognitive science. *Mind, Brain, and Education, 9*, 100–106.
Venkadasalam, V. P., & Ganea, P. A. (2018). Do objects of different weight fall at the same time? Updating naive beliefs about free-falling objects from fictional and informational books in young children. *Journal of Cognition and Development, 19*, 165–181.
Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition, and analogical reasoning in older adults. *Psychology and Aging, 19*, 581–591.
Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*, 375–382.
Walker, C. M., Gopnik, A., & Ganea, P. A. (2015). Learning to learn from stories: Children's developing sensitivity to the causal structure of fictional worlds. *Child Development, 86*, 310–318.

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28*, 1205–1212.
Waxman, S. R., Herrmann, P., Woodring, J., & Medin, D. (2014). Humans (really) are animals: Picture-book reading influences 5-year-old urban children's construal of the relation between humans and non-human animals. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00172.
Waxman, S. R., Lynch, E. B., Casey, K., & Baer, L. (1997). Setters and Samoyeds: The emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology, 33*, 1074–1090.