Acta Psychologica 224 (2022) 103505

Contents lists available at ScienceDirect

Acta Psychologica

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

Why do children struggle on analogical reasoning tasks? Considering the role of problem format by measuring visual attention☆

Katharine F. Guarino[a,*], Elizabeth M. Wakefield[a], Robert G. Morrison[a], Lindsey E. Richland[b]

[a] Loyola University Chicago, United States of America
[b] University of California, Irvine, United States of America

ARTICLE INFO

Keywords:
Analogy
Attention
Eye tracking
Reasoning
Pattern recognition
Children

ABSTRACT

Given the importance of analogical reasoning to bootstrapping children's understanding of the world, why is this ability so challenging for children? Two common sources of error have been implicated: 1) children's inability to prioritize relational information during initial problem solving; 2) children's inability to disengage from salient distractors. Here, we use eye tracking to examine children and adults' looking patterns when solving scene analogies, finding that children and adults attended differently to distractors, and that this attention predicted performance. These results provide the most direct evidence to date that feature based distraction is an important way children and adults differ during early analogical reasoning. In contrast to recent work using propositional analogies, we find no differences in children and adults' prioritization of relational information during problem solving, and while there are some differences in general attentional strategies across age groups, neither prioritization of relational information nor attentional strategy predict successful problem solving. Together, our results suggest that analogy problem format should be taken into account when considering developmental factors in children's analogical reasoning.

1. Introduction

Analogical reasoning involves identifying higher order similarities in relational structure shared between representations (Gentner, 1983, 2010). This form of reasoning is used in many contexts and is a core part of expertise in many professional disciplines (see Gentner et al., 2001; Richland & Simms, 2015). In its mature form, it is a powerful cognitive mechanism that contributes to a range of skills unique to humans, including innovation (Markman & Wood, 2009), creativity (Holyoak & Thagard, 1995), inductive reasoning (Holland et al., 1986), general intelligence (Snow et al., 1984), and even cultural evolution (Brand et al., 2021; for review see Holyoak, 2012). Each of these skills require identifying relational similarity shared between contexts in order to make a conclusion about the contexts as a whole.

However, analogical reasoning proficiency develops over time: Young children often struggle to notice or extract deep underlying structures from comparison opportunities (e.g., Gentner, 1988; Holyoak et al., 1984; Rattermann & Gentner, 1998; Richland et al., 2006; Thibaut et al., 2010; Thibaut & French, 2016). One task that is often used to test analogical reasoning ability is a propositional analogy in the format A:B::C:D. In these analogies, participants select from four choices a D item that is relationally similar to the C item in the same way that A and B are similar. For example, in a pictorial propositional analogy, item A might be a picture of a peanut, item B a picture of an elephant, and item C a picture of a banana. The relation between items A and B is that elephants *like to eat* peanuts, so the goal is to identify which option from an array is something that *likes to eat* bananas. In this example, four options that are similar to the banana in different ways would be presented in the choice array: a monkey (relationally similar), a yellow canoe (featurally similar), an orange (semantically similar), and a wrench (neutral or unrelated). The correct choice would be the monkey, because monkeys *like to eat* bananas. However, when asked to solve these problems children struggle to identify the relation shared between the A and B items to find a missing D item from an array – they struggle to extract the underlying relational structure of the comparison.

Given the importance of developing this ability, researchers have asked why reasoning analogically is so challenging for children. Two common explanations in the literature implicate: 1) children's inability

☆ The data that support the findings of this study are openly available in Open Source Framework at https://osf.io/ypra8/?view_only=1f51ed66f9ac446fad94917a6497f7b9, DOI:10.17605/OSF.IO/YPRA8.

* Corresponding author at: Department of Psychology, Loyola University Chicago, 1032 West Sheridan Rd, Chicago, IL 60660, United States of America.
E-mail address: kguarino@luc.edu (K.F. Guarino).

https://doi.org/10.1016/j.actpsy.2022.103505
Received 19 June 2020; Received in revised form 26 September 2021; Accepted 14 January 2022
Available online 25 January 2022
0001-6918/

to prioritize attending to relational information (the A:B relation in a propositional analogy, e.g., the peanut:elephant relation) during initial problem solving, or 2) children's inability to disengage from featurally salient distractors (e.g., the yellow canoe, which is featurally similar to the banana). While there is support for both explanations, recent work with propositional analogies suggests that the root cause of children's difficulty is the former explanation: their lack of prioritization of relational information (Starr et al., 2018).

The purpose of the present study is to examine whether this conclusion by Starr et al. (2018) – that failure to prioritize relational information is the core reason children struggle with propositional analogy problems – is generalizable across analogy formats. Previous literature exploring the impact of problem format in other domains has found that the appearance of problems impacts children's apparent skills (e.g., numerical cognition: Cantlon et al., 2007; Leung & Silver, 1997; Mix, 1999; Siegel, 1974, object matching tasks: Paik & Mix, 2006, word recognition: Walley et al., 1995), such that featural task components are difficult for children to ignore during problem solving (e.g., Mix, 1999; Paik & Mix, 2006; Siegel, 1974). As one example, Cantlon et al. (2007) found that when asking 3–5 year-olds to estimate the number of items in an array, their performance was impaired by visual heterogeneity among the items, such that they were significantly better at estimating when the items had similar sizes, shapes, and color elements. With this body of work in mind, we predict that children's approach to problems of analogy will also be impacted by problem features and format, and thus, the conclusions drawn by Starr et al. (2018) may reflect characteristics of the propositional analogy paradigm, and may not generalize to analogical thinking in other task formats perhaps more similar to the way analogy is used in everyday thinking.

For the purpose of comparison, we consider another well studied analogy problem type: scene analogies. When solving scene analogies, participants are asked to map role correspondences between two scenes (e.g., a source and target scene) that both depict the same relation. For example, a pair of scenes might depict a dog chasing a cat (source relation) and a boy chasing a girl (target relation). The experimenter highlights one object in the source scene (e.g., the cat - the patient of the source "chase" relation) and asks participants to select the corresponding part of the same relation in the target scene. The correct choice would be the girl, because cat and girl are both *being chased*. These analogies can also introduce a featural match, often referred to in the literature as a featural distractor – in this case, a cat placed in the target scene. Just as in a propositional analogy, children may incorrectly solve this type of problem because of 1) an inability to prioritize relational information (the items in a relation within the source scene, in this example, the dog chasing a cat) or 2) an inability to disengage from items in the target that are featurally similar to the indicated item in the source scene (the cat in the target scene which is featurally similar to the cat in the source scene).

1.1. Children have difficulty prioritizing relational information during analogical reasoning

One explanation for children's difficulty with analogical reasoning is that they fail to notice the presence or relevance of relational correspondences, even when it is key to solving an intended task or problem-solving paradigm. From eye tracking work, we know when adults solve propositional analogies, they generally attend to the A:B pair before fixating on C and the response choices, showing that they maintain the overarching goal when solving the problem (i.e., find the picture that goes with C in the same way that A goes with B; Glady et al., 2016; Starr et al., 2018). Using our previous propositional analogy example, this would mean adults look to the elephant and peanut, presumably to ascertain how the items are related, before considering the banana and possible response options. Similar conclusions about adults' visual attention during analogy solving have been demonstrated when using scene analogies. When asked to solve problems of this format, adults primarily attend to the source relation (i.e., the dog chasing the cat in the previous example) before exploring the target scene and response options (Glady et al., 2016; Gordon & Moser, 2007). Therefore, across analogy formats, it seems that a mature visual attention pattern is characterized by prioritizing the source pair of items (A:B pair or items involved in the source relation) and identifying the relation between the items before applying that relation to the target space (see also Sternberg, 1977).

In contrast, 5- and 6-year-old children tend to ignore the A:B pair when solving propositional analogies, and focus their attention on C and the response choices (Glady et al., 2017; Thibaut & French, 2016). Referring to the previous propositional analogy example, this would mean children look to the banana and response options without identifying the underlying relation for which they are solving. The implications of this attentional difference suggest that children are less likely than adults to extract relational information before considering response options, instead, focusing on the immediate task goal (i.e., find the picture that goes with C). This difference in looking patterns has consequences: Using linear discriminant analysis, French and Thibaut (2014) demonstrated that children's visual attention during the first third of a trial predicted with 64% accuracy whether or not the problem was answered correctly. The more children initially looked to the A:B pair, the more likely they were to answer the trial correctly, and the more they ignored the A:B pair, the more likely they were to answer incorrectly. Glady et al. (2017) provide further support for these findings by demonstrating that guiding children's attention to the A:B pair during initial problem solving significantly improved their performance. Similarly, Chen et al. (2016) found that when children are asked to solve three-by-three matrix completion problems, which require similar abilities as an analogy problem, they struggle to utilize the relational problem structure until they are prompted to engage in effective solving strategies. Children's performance improved significantly when effective visual attention strategies (i.e., attending to relational problem structure) are encouraged early on during solving (Chen et al., 2016; see also Hosenfeld et al., 1997).

1.2. Children are drawn to featurally salient distractors during analogical reasoning

Although attending to relational correspondences is important for analogical reasoning problem solving, another line of research suggests that children struggle with these problems because of featurally salient distractors. In many situations that require analogical reasoning, the visual scene is complex. Although higher-order relational or structural correspondences are present, children are more likely to make judgments based on surface-level similarities – attending to items that are featurally related to the item in question, rather than structurally related. Using the previous example, children are more likely to choose the yellow canoe as being related to the banana (C item) because these items are featurally similar, even though the monkey is the correct, relational choice. Young children are particularly susceptible to this type of error, tending to shift from more object-based similarity matching to more relational reasoning over time, defined as the *relational shift* (Gentner, 1988).

This shift has been posited to stem from knowledge accretion (e.g., Rattermann & Gentner, 1998), but growing literature suggests that even when knowledge is adequate, children still show difficulty inhibiting attention to objects that draw attention away from relations (Richland et al., 2006; Thibaut et al., 2010). For example, when 3–4 year old children were asked to solve scene analogy problems that contained a featural distractor, an item in the target scene that shared surface similarity to the prompted item in a source scene but did not relationally correspond, the perceptually similar match was an effective lure. Specifically, children's accuracy for problems with distractors (e.g., a cat in the target scene that is perceptually similar to the prompted cat in the source scene) was 15% lower than that for problems without distractors

(Richland et al., 2006). These behavioral findings have been complemented by modeling work: Simulations in the MAC/FAC model of analogy (Many Are Called/Few Are Chosen; Forbus et al., 1995) suggest initial solving is characterized by similarity-based retrieval processes. Similarly, simulations in the LISA computational model of analogy (Learning and Inference with Schemas and Analogies, Hummel & Holyoak, 1997, 2003) suggest that changes in inhibition levels, along with relational knowledge accretion, account for young children's difficulty when reasoning analogically (Doumas et al., 2018; Morrison et al., 2011). Using a scene analogy task, LISA replicates the experimental findings of Richland et al. (2006), such that the model was more likely to choose a featurally similar distractor object than an analogically correct choice.

Thibaut and French (2016) provide additional evidence that correct responses on analogy tasks may be tied to visual attention, such that relational reasoning in children is correlated with directing visual attention away from salient, yet irrelevant, featural corresponding objects and towards the key relational correspondences. Replicating the scene analogy data, children's performance on a propositional analogy task correlated negatively with presence of a perceptually salient distractor (Thibaut et al., 2010). Further, eye tracking revealed a negative association between the amount of time looking to a distractor and performance, such that the more time children spent looking at the distractor the worse they performed (Thibaut & French, 2016).

1.3. Failure to prioritize relations versus focus on featural distractors

Thus, there is reason to believe that the children's typically low performance on analogy tasks is due to either a lack of ability to notice and abstract key relational correspondences from a source context (i.e., to prioritize relational information) or difficulty ignoring salient perceptual similarity (i.e., to disengage from salient distractors). Whereas much of the previous literature has tested the viability of one or the other explanation for children's performance, recent work has considered both sources of error within the same study to determine whether one has more explanatory power than the other. Using propositional analogies, Starr et al. (2018) pitted these two explanations against each other and used eye tracking methodologies to gain direct data on attentional patterns. By categorizing children's visual attention patterns as either prioritizing relational information within a problem (i.e., the A:B relation) *or* focusing on the target relation (i.e., the C item) and response options, as noted above they found that children, compared to adults, were more likely to ignore the A:B pair and focus their attention on C and response options. This visual attention pattern was associated with choosing the response option that was featurally similar to the C item. Starr and colleagues also found that initial looking patterns, rather than attention across the course of each problem, strongly predicted children's likelihood of selecting the correct, relational response. Specifically, children's initial focus on the relational information contained in the A:B pair was positively related to their performance, suggesting that focusing on relational information from the start is necessary for successful analogical reasoning. The authors therefore concluded that poor performance was due to an inability to prioritize attending to the A:B relation when initially processing an analogy, *rather* than an inability to disengage from featural lures.

Thibaut and French (2016) assessed the proportion of time children and adults spent looking towards each item within A:B::C:D analogies, and suggest that without an initial focus on A and B items, children are less able to inhibit a focus on items that are featurally similar to the C item. Together, this work supports the argument that the root of children's difficulties in solving analogies comes from an inability to initially focus on source relations. However, what is unknown is whether this finding has to do with the characteristics of propositional analogy problems and the constraints of finding key relations within that context, or whether it is a generalizable characteristic of children's analogical problem solving across other forms of representing analogy.

1.4. Current study: does problem format matter?

In the present study, we used eye tracking to monitor visual attention while children and adults solved scene analogy problems similar to those used by Richland et al. (2006). Scene analogy problems are informative because they are more ecologically similar to how children might make visual analogies between role relations in everyday settings than propositional analogies. Additionally, they avoid the problem that propositional analogy problems can be solved without analogical mapping, but rather can be solved as an analogy verification where the reasoner just decides whether the source and target exhibit the same relation. In four term propositional analogies one can abstract a relation from the A:B representation, e.g. "mother of", and then in the target C:D relation, they can solve the problem by simply verifying that D is in fact a mother of C ("true or false"). This allows the reasoner to solve this problem in two parts, such that neither requires higher order relations between the A:B and C:D representations. In a scene analogy, in contrast, in order to identify the correct object in the target problem (e.g. the chaser in a chase relation), one must necessarily look for both an agent and a patient of a chase relation out of a range of options to determine that this is indeed depicting chasing (e.g. not simply a person running), and they will need to ensure that the selected object's role is in common with the identified source object (chaser vs. chase). Thus, this problem format can be a particularly informative test of children's attentional and error patterns.

We focus on how children and adults attended to source relation items and featural distractors within the scene analogies, and whether visual attention predicted behavioral performance. This allowed us to assess which of these factors best explains the difficulty children are known to display when solving scene analogies. If the main factor underlying children's poor performance on scene analogy problems is a strategy failure of not prioritizing attention to relational structures, we should find that adults allocate a greater amount of their visual attention to the source scene and key relationship than children, especially early in problem solving. We should also find that visually attending to the source relation is positively related to behavioral accuracy on problems. These findings would suggest that across analogy formats, children's lack of priority for spontaneous attention to structured relations is a root cause of their failure to successfully solve analogy problems. However, if the difficulty with solving scene analogies for children lies in their inability to disengage from featural distractors, we should find that children allocate a greater amount of their visual attention to the featural distractor in the target scene than adults, and that this is negatively related to behavioral accuracy on problems. This finding would suggest that problem format affects how children approach an analogical reasoning problem, and that different analogy formats impact children's use of strategy deployment for solving analogy problems.

Beyond comparing how these two strategies explain visual attention allocation during reasoning, we also ask *how* children solve analogy problems and if problem format impacts their strategy use. Previous work has used a theoretically motivated approach to ask this question when assessing children and adults' strategies for solving propositional analogies. While many models have been put forth to explain the processes involved in analogical reasoning, three classic models have been the primary focus of work examining visual attention during analogy solving: projection-first, structure-mapping, and semantic-constraint. In the present study we consider children and adults' looking patterns in relation to these three models of analogy problem solving. These models vary in the emphasis placed on assessing the source relation before resolving the target relation, comparing across the source and target relations, and focusing predominately on the target relation and response options, respectively (e.g., Chalmers et al., 1992; Doumas et al., 2008; Falkenhainer et al., 1989).

The first theory put forth, and what many consider the optimal strategy to engage in during analogical reasoning, is a projection-first strategy (e.g., Doumas et al., 2008; Hummel & Holyoak, 1997;

Sternberg, 1977). In this model, propositional analogies are solved by identifying the rule that relates the A and B items, and then applying that rule to the C item to determine the solution item D. This strategy emphasizes generating the relation between A and B in order to systematically constrain one's attention in the search space when trying to find a match for C. This strategy closely mirrors adults' visual attention patterns during analogy solving (Glady et al., 2016; Gordon & Moser, 2007; Starr et al., 2018; Thibaut & French, 2016).

The second theory emphasizes alignment between the A and C items and between the B item and the target, rather than attending to the source relation (A:B) initially, and is therefore, called a structure-mapping model (Falkenhainer et al., 1989; Gentner, 1983, 2010). In this model, analogies are solved by mapping items from one structured representation to another. Using the previous example, this would require initially generating a rule to relate the peanut and the banana (e.g., both something that can be eaten), and then identifying a D item (the monkey) that relates to the elephant (e.g., both animals). While a projection-first strategy employs a top-down matching approach by establishing a rule between the A and B items initially, a structure-mapping strategy employs a more bottom-up approach by generating possible sets of correspondences between A and C and choosing one that fits with B.

The final model that is discussed in the literature as a common approach to solving analogies, especially for young children who have difficulty extracting relational or structural information from comparisons, is a semantic-constraint strategy (Chalmers et al., 1992; Thibaut & French, 2016). This model is based on the assumption that many possible relations exist between the A and B items, and that attention should be focused on the C item and the response options in order to narrow the hypothesis space. While this could be a logical approach, it is not as efficient as generating an initial, guiding rule between A and B items (projection-first) or identifying the relational structure shared across contexts (structure-mapping). This early focus on the C item and response options constrains the search space without prioritizing either the A or B items.

In general, previous eye tracking work using propositional analogies suggests that adults are most likely to utilize a projection-first strategy, rather than a semantic-constraint or structure-mapping strategy, and that using this strategy is most likely to lead to a correct response (Thibaut & French, 2016; Vendetti et al., 2017). Similar conclusions have been made when using scene analogies (Gordon & Moser, 2007), such that when solving scene analogies adults process source relations and generate rules between objects before attempting to apply that rule to another context. In contrast, children are more likely to utilize a less efficient semantic-constraint strategy (Starr et al., 2018). When comparing children and adults' visual attention when solving a variety of scene analogy problem types, Glady et al. (2010) found that overall, children's scanpaths are more variable than adults' scanpaths, suggesting that adults utilize a more consistent strategy to solve propositional analogy problems. And while there is less evidence that children or adults' visual attention patterns align with a structure-mapping strategy, there is evidence that other types of analogies utilize structure-mapping (e.g., Gentner, 1983; Markman & Gentner, 1993). For the first time, we ask how these three models can be used to understand the approach adults and children take when solving scene analogy problems, and whether these strategies can be used to explain success or failure on analogy problems.

2. Method

2.1. Participants

Data from 57 4- and 5-year-old children[1] (29 females, $M_{age} = 4.88$, $SD_{age} = 0.47$) and 45 adults (37 females, $M_{age} = 19.45$, $SD_{age} = 0.99$) were analyzed for the present study.[2] Participants represented a diverse sample from a large metropolitan city (Children: 48% White, 14% Black, 5% More than one race, 5% Asian, 28% Unreported; Adults: 63% White, 4% Black, 3% More than one race, 12% Asian, 18% Unreported). Children were recruited from schools and participated individually in one experimental session during a regular school day. Children were compensated with stickers and classroom teachers were provided with a gift card for classroom materials. Adults were recruited from a university participant pool and participated individually in a lab. They received course credit for their participation. All participants and/or guardians completed informed consent and/or assent procedures and were treated ethically.

2.2. Materials

2.2.1. Stimuli

Participants were shown scene analogies adapted from Richland et al. (2006). Each stimulus included a pair of scenes presented simultaneously on a *15-inch* Dell laptop. Pairs of scenes depicted one of two *relation categories* (i.e., chasing or reading[3]) occurring between items (i.e., animals or people). Source scenes contained five items: two items within the relation, and three additional items (i.e., neutral inanimate objects). Target scenes also contained 5 items: two items within the relation, two additional items, and a featural distractor.

Fig. 1A shows a "*chasing*" relation. The source scene on the left shows a tiger *chasing* a woman (items within the chasing relation), and the target scene on the right side shows a lion *chasing* a horse (items within the chasing relation). Target scenes also contained a featural distractor

[1] Children of this age have the potential to succeed on analogical reasoning tasks when they have adequate domain knowledge, yet struggle when featural distractors are present and external support is not provided (Richland et al., 2006). Therefore, these children are at the ideal point in development to assess the effects of instruction.

[2] Data from 57 children and 60 adults was collected. Although all children were included in analyses, a subset of data from particular timepoints were excluded from 8 children based on insufficient usable eye tracking data. Five adult participants were excluded for having lacking sufficient eye tracking data. For adult participants to be included in the sample, they must have >75% accuracy. This was to ensure that we had a measure of successful, mature visual attention patterns. Ten adult participants were excluded for having <75% accuracy across trials.

[3] The current study was part of a larger study, where the aim was to assess whether learning strategies differ across items. The block of trials used here were the pretest of this larger study, where participants were trained using the chasing relation. In the larger study the experimental manipulation presented participants with verbal instruction explaining the relational structure of four scene analogies either with or without accompanying hand gestures that highlighted the relations across and within scenes. The relations of *chasing* and *reading* were chosen because they are best reflected through gesture. Following training we could test not only learning at posttest on chasing relations, but also the ability of instruction to support generalization of solving strategies to reading relations, which participants were not trained on. While a comprehensive discussion and analysis of comparisons between trial types is beyond the scope of the present study, briefly we explored whether children's performance was impacted by trial type (chasing vs reading) given the hypothesis that problem format may play a larger role in children's solving strategies than adults. However, we did not find a significant difference between trial types ($\beta = 0.04$, $SE = 0.03$, $t = 1.05$, $p = .29$). The impact of trial type should be further explored in future work to take a closer look at whether children's solving strategies differ by scene analogy type.

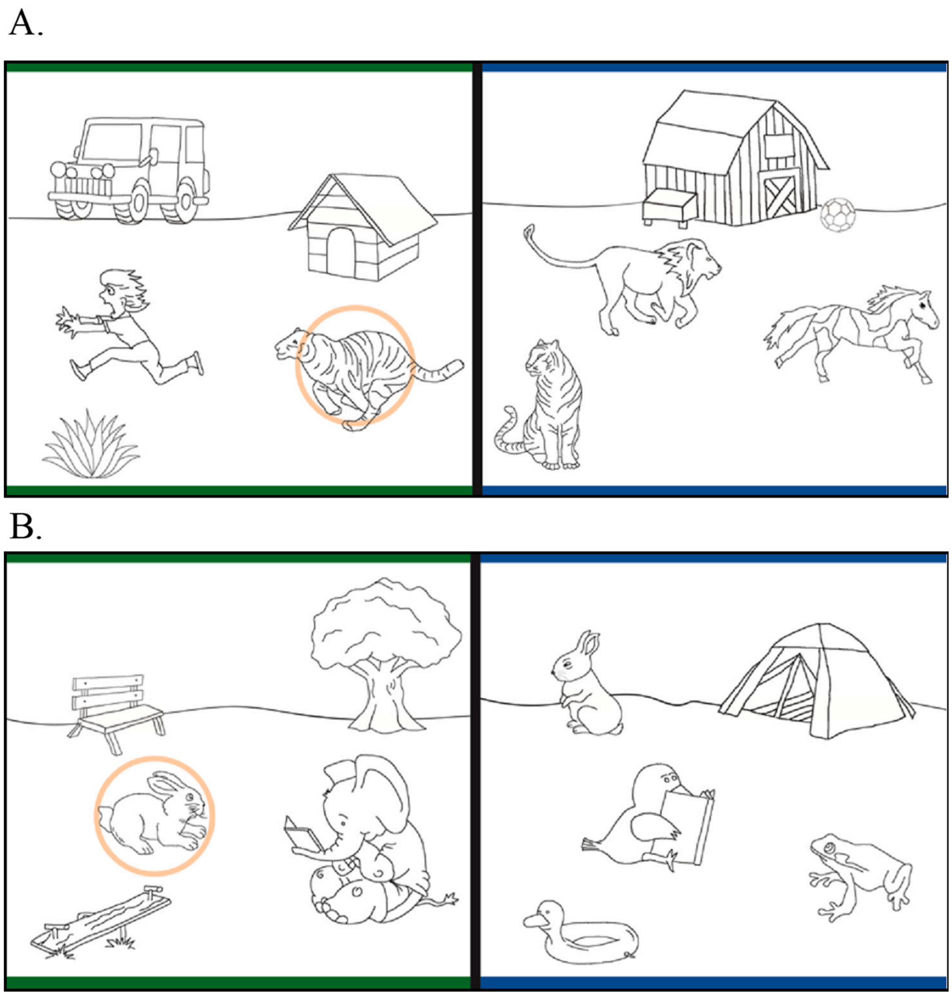

Fig. 1. Source scene (left) and target scene (right). A. Example trial of *chasing* relation category. B. Example trial of *reading* relation category.

that was featurally similar to the prompted source-scene item; the tiger in the target scene serves as the distractor because the tiger is prompted (i.e., circled) in the source scene. Similar prompts have been used in previous work (e.g., arrows in Richland et al., 2006). Distractors were centrally located within the stimuli (across both pair of scenes), increasing the likelihood that participants would notice them. To maintain the same number of items across scenes, *neutral* items were included (Fig. 1A: dog-house, jeep, plant, barn, soccer ball).

Fig. 1B shows a "*reading*" relation. Items depicting the reading relation were oriented towards each other with one character reading to the other character. In all source scenes, one of the two items within the source relation was prompted with a circle. Directionality of relations within a pair of scenes was reversed to avoid children making choices based on spatial location alone.

2.2.2. Eye tracker

Eye tracking data were collected via corneal reflection using a TobiiPro X3-120 remote eye tracker mounted at the base of a 15-inch Dell laptop screen. Tobii software was used to perform a 5-point calibration procedure. Then, gaze data were collected and integrated with video stimuli (described below) using Tobii Studio (Tobii Technology, Sweden). Data were extracted on the level of individual fixations and fixation location was queried at 8.33 ms intervals using R studio, to align with the native sampling frequency of the eye tracker.

2.3. Procedure

For the present study, we analyzed eye tracking data collected while participants solved scene analogy problems. These data were drawn from a longer study. For children, the data came from 12 pretest problems (6 chasing; 6 reading); in the longer study, children then received instruction on how to solve scene analogies and completed 12 posttest problems. For adults, the data came from 24 problems (12 chasing; 12 reading). Items included in a child's pretest and posttest were counterbalanced, and all items were shown to adults.

2.3.1. Introduction to task and calibration

Participants were told they were going to play a picture game and shown one example trial, orienting them to the layout of test trials (i.e., two pictures with different colored borders) and their task (i.e., for each set of scenes their job was to find the pattern in the pictures). The experimenter described the chasing relation and asked the participant to solve the relation. For children, the explanation was repeated until they chose the correct item. This introduction ensured that when children incorrectly answered a trial, it was not because they misunderstood the goal of the task. Next, calibration of the eye tracker was completed: Participants were seated approximately 40 cm in front of the laptop and familiarized with the eye tracker.

2.3.2. Task

Participants completed a set of scene analogies while their visual attention was monitored. Participants were instructed to respond verbally to question "*Which thing in the picture with the blue edges is in the same part of the pattern as the circled thing in the picture with the green edges?*" The task was self-paced, but if no response was given after a few seconds, the experimenter re-prompted participants. Responses were recorded.

2.4. Data analysis

2.4.1. Proportion of looking to AOIs

Areas of interest (AOIs) were created for the items within the scene pairs using Tobii Studio (i.e., each trial had 10 AOIs, 5 in each scene).[4] The remaining spaces outside of these AOIs were collapsed into an "Other" AOI. For analyses, we considered visual attention to 1) the source relation (comprised of two relational items and analogous to the A:B pair in propositional analogies) and 2) the featural distractor (analogous to the featuralchoice item in propositional analogies). Proportion of time spent looking to each AOI was calculated by dividing the amount of time spent looking at a given AOI during a trial by the total time looking at the screen during that trial. All analyses considered visual attention patterns and accuracy at the trial level.[5]

Because prior work (French & Thibaut, 2014; Starr et al., 2018) suggests visual attention during initial solving has the most predictive power for whether a participant will arrive at the correct answer, we consider proportion of looking to these AOIs across the entire trial, as well as proportion of looking during the first third of each trial, as this method has been used to assess initial looking patterns in adults and children (French & Thibaut, 2014; Thibaut & French, 2016). To comprehensively understand how solving strategies differ between children and adults we examined visual attention patterns across all trials, correct and incorrect, to assess how looking patterns relate to performance. This analytic approach has been used previously to characterize looking patterns associated with successful analogical reasoning problem solving. For example, Thibaut and French (2016) have examined the relation between children's performance and visual attention patterns on propositional analogies to characterize what visual strategies are associated with accuracy and errors. While their primary analysis was to characterize looking patterns during correct trials, they performed post-hoc analyses to compare how eye-tracking gaze patterns differ between correct and incorrect trials.

2.4.2. Visual attention pattern classification

In addition to considering proportion of time spent looking to individual AOIs, we performed classification analyses on participants' visual attention patterns during problem solving. We considered three classic models that have been put forth to explain the processes underlying analogy problem solving: projection-first, structure-mapping, and semantic-constraint models. Following the procedures used by Starr et al. (2018) and Vendetti et al. (2017), classification algorithms were developed for each model, informed by the assumptions of each associated theory and adapted to suit the format of scene analogies.

Some aspects of propositional and scene analogies are comparable: Just like in a propositional analogy, it is necessary to identify the primary relation between items in a scene analogy. That is, the solver needs to identify the relation represented in the source scene, and what items are involved in that relation (W and X in Fig. 2). Similarly, both problem formats contain a featural distractor as a response option (X′ in Fig. 2), such that the C item in a propositional analogy and the X item in a scene analogy both perceptually match one of the response options.[6]

However, while some aspects of propositional and scene analogies are easy to align (e.g., the functionality of the featural distractors), there are some notable differences in these formats that need to be carefully considered when adapting the classification algorithms used in previous literature for use in the present study. For example, while the A:B items in a propositional analogy are treated as the 'source relation', and the C item is the 'prompted' item that needs to be matched with the target response option, in a scene analogy, there is no true comparison to the C item. In a scene analogy, the source relation in the source scene *contains* the prompted item that needs to be matched to an item in the target scene. Therefore, in a scene analogy, the prompted item is both *incorporated* in the source relation and is what needs to be *matched* to the target response option. In a propositional analogy it is advantageous to identify the relation between the A:B items and then map that relation to the C item, but in a scene analogy, you need to both identify the relation within the source scene *and* map relations across scenes (e.g., two items are chasing, two items are being chased). Further, in a scene analogy there are also more response options (5 items in the target scene) vs the 4 items in a propositional analogy. We adapted the algorithms from previous work with these issues in mind.

Broadly, the classification algorithms generate a score that accounts for the order in which an individual fixates on items in the scene analogy and the amount of time spent looking at individual items. The higher the score, the better the fit between the visual attention pattern during a trial and the model. Table 1 provides classification algorithms used for propositional analogies (Vendetti et al., 2017) vs scene analogies (developed for the current study). The order of the items listed in the classification models is not related to how they are calculated. Rather it is the value of each item in the models that differs depending on the sequence of visual attention, where fixations are weighted such that they receive higher scores (or values) in the models if they occur earlier on within the participants' sequence of fixations and transitions.

According to the projection-first model, initial fixations and transitions occur within the source domain. In a propositional analogy, this strategy is characterized by initial fixations to and transitions *between* the A and B items. Likewise, in a scene analogy, this strategy is characterized by fixations and transitions *within* the source scene – specifically, between the two items involved in the relation in the scene on the left (items W and X). In a scene analogy, these initial fixations and transitions are followed by attempts to map the relation in the source scene to the target scene by identifying which item in the target scene serves the same role in the relation as the prompted item. According to Starr et al. (2018), this initial looking pattern is the most advantageous for extracting the relation that is shared between scenes, and identifying the correct, relational choice in the target scene.

According to the structure-mapping model, initial looking patterns include fixations to and between corresponding relational items in the source and target scenes (see Table 1). In a propositional analogy, this corresponds to fixations and transitions between the A and C items and B and D items, and in a scene analogy, this corresponds to fixations and transitions between the two items chasing (W and Y) and the two items being chased (X and Z) (see Fig. 2). In a scene analogy, this pattern of

[4] AOIs were drawn following the contours of the individual items with a slightly extended border to account for possible precision issues with calibration.

[5] Analysis of aggregated data demonstrated the same pattern of findings as described in the following sections using trial level data.

[6] For semantic-constraint strategy, we had to consider which item within the scene relation was prompted, and therefore, which item in the target relation is the correct choice. For half the trials, the item *chasing* was prompted and for the other half the item *being chased* was prompted. For trials in which the item *being chased* is the prompted item and, therefore, the correct choice, we reassigned the AOI labels, such that for all trials, labels X and Z are associated with the items *chasing*, and labels W and Y are associated with the items *being chased*. This allowed us to make comparisons across all trials, despite *which* item in the relation is prompted.

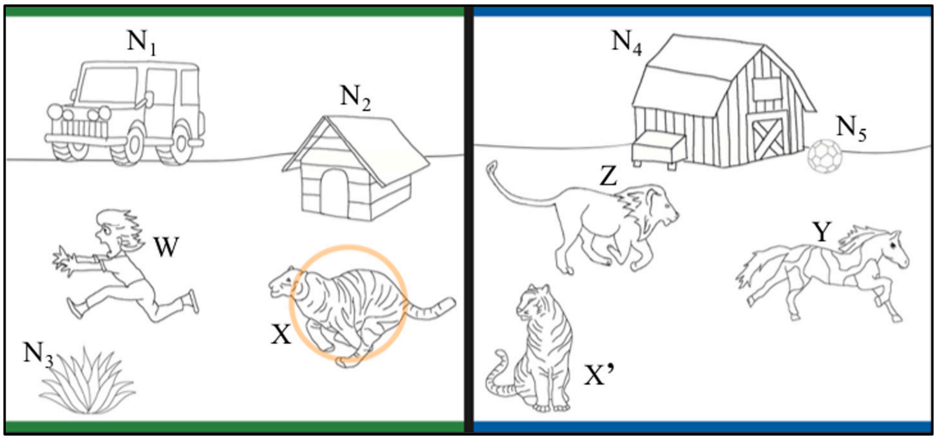

Fig. 2. Example scene analogy problem with items labeled, with X′ indicating the distractor that is featurally similar to the prompted item X.

Table 1
Classification algorithms for propositional and scene analogies.

Strategy	Propositional analogies	Scene analogies
Projection-First	$Score_{initial\ fix\ A} + Score_{initial\ fix\ B} + Score_{AB} + Score_{BA} + Score_{CT}$	$Score_{initial\ fix\ W} + Score_{initial\ fix\ X} + Score_{WX} + Score_{XW} + Score_{YZ} + Score_{ZY}$
Structure Mapping	$Score_{initial\ fix\ A} + Score_{initial\ fix\ C} + Score_{AC} + Score_{CA} + Score_{BT}$	$Score_{initial\ fix\ W} + Score_{initial\ fix\ Y} + Score_{initial\ fix\ X} + Score_{initial\ fix\ Z} + Score_{WY} + Score_{YW} + Score_{XZ} + Score_{ZX}$
Semantic-Constraint	$Score_{initial\ fix\ C} + Score_{initial\ fix\ T} + Score_{initial\ fix\ S} + Score_{initial\ fix\ P} + Score_{initial\ fix\ U} + Score_{CS} + Score_{SC} + Score_{CP} + Score_{PC} + Score_{CT} + Score_{TC}$	$Score_{initial\ fix\ Y} + Score_{initial\ fix\ Z} + Score_{initial\ fix\ D} + Score_{initial\ fix\ N4} + Score_{initial\ fix\ N5} + Score_{XY} + Score_{YX} + Score_{XZ} + Score_{ZX} + Score_{XD} + Score_{DX} + Score_{XN4} + Score_{N4X} + Score_{XN5} + Score_{N5X}$

Note. In propositional analogy equations, T indicates the target (the correct response), S indicates the response option that is a semantically related to the C term, P indicates a response option that is a perceptually related to the C term (i.e., similar to the featural distractor in scene analogies), and U indicates a response option that is unrelated both perceptually and semantically to the C term. In scene analogy equations, X indicates the prompted item within the chasing relation in the scene on the left, W indicates the other item in the chasing relation in the scene on the left, Z indicates the item corresponding to X in the scene on the right, Y indicates the item corresponding to the W in the scene on the right, X′ indicates the featural distractor, and $N_{1–5}$ indicate the additional items in the scenes that are not involved in the chasing relations. In each equation, ***initial fix*** corresponds to the first instance a participant fixated on the specified AOI (i.e., not necessarily the first fixation within a given sequence of fixations). Scores based on two items (e.g., WX, CP) represent transitions between these items. All transitions are reciprocal (e.g., $Score_{WX}$ and $Score_{XW}$ account for a reciprocal transition between AOIs for items X and W).

mapping across scenes should then lead to the identification of the item that serves the same role as the prompted item, given that this cross-scene relation is identified initially.

Finally, according to the semantic-constraint model, visual attention is initially oriented to the target relation and the possible response options (see Table 1). Most work that has considered this model has done so within the format of a propositional analogy, where initial attention is oriented towards the C item and response options. Because there is not a true equivalent component in a scene analogy to the C item, we adapted the equation to suit the scene analogy format. In a propositional analogy, the immediate task goal is 'find the item that goes with C', and the strategy is to look at the response options and compare them to this goal. In a scene analogy, the immediate goal is 'find the item in the target scene that goes with the circled item', and the response options are the items in the target scene. Thus, we considered a semantic-constraint strategy to be characterized by an initial focus on the target scene items (the response options: X′, Y, Z, N_4, and N_5; see Fig. 2) and transitions between those items and the prompted item (X). Specifically, we consider initial fixations to all response options, given that when utilizing a semantic-constraint strategy with propositional analogies, solvers jump to matching the C item and the response options, and in a scene analogy, it is likely that the solver will also jump to the target scene response options to find 'what in the *target scene* goes with the circled item'.

Scores were calculated using fixation duration and the ordinal position of a particular fixation or transition within a participants' sequence of fixations for a given trial. Each score was calculated as the product of its duration and inverse of its location within the fixation sequence for a given trial. Score value assumes that fixations and transitions that occur earlier in the trial and those of longer duration are indicative of the strategy being used. Therefore, fixations and transitions receive higher scores the earlier and longer they occur. Scores for individual strategies were calculated by summing the individual fixation and transition scores for the given equation. Trials were classified as the highest-scoring strategy. For example, a trial would be classified as projection-first if the score for this strategy was higher than for semantic-constraint or structure-mapping. To remain consistent with classification procedures used by Starr et al. (2018) and Vendetti et al. (2017), in cases where a trial had a tie between strategies scores, that trial was excluded from the respective analyses.

3. Results

First, we considered broad differences between how children and adults solved scene analogy problems. On average, children made 36.05 fixations per trial (SD = 8.27) and adults made 18.93 fixations per trial (SD = 5.06), and adults performed the task more quickly than children (adults: M = 5220 ms SD = 1240 ms; children: M = 13,130 ms; SD = 3330 ms). These differences reached statistical significance (fixations: $\beta = 16.54$, $SE = 1.63$, $t = 10.16$, $p < .001$; problem-solving time: $\beta = 7845.04$, $SE = 536.37$, $t = 14.63$, $p < .001$), suggesting that children spend more time solving these problems and make more fixations per trial before determining their answer than adults. Over the course of the task, children did show a decrease in problem-solving time ($\beta = -534.31$, $SE = 59.07$, $t = -9.05$, $p < .001$), but accuracy was not predicted by trial ($\beta = -0.00$, $SE = 0.03$, $t = -0.20$, $p = .85$) and solution time did not predict accuracy ($\beta = -0.00$, $SE = 0.00$, $t = -0.70$, $p = .48$), suggesting that no matter the reason for children's increased reaction times, this did not affect their performance.

3.1. Prioritization of relational information

Our primary goal was to establish whether visual attention to the source relation during scene analogy problems differed between age groups in the same way as for propositional analogies. Fig. 3 shows the proportion of visual attention allocated to AOIs for both children and adults. In contrast to previous work (e.g., Starr et al., 2018) using propositional analogies demonstrating that adults spend more time focusing on the source relations than children, both groups attended to the source relation about one-third of the time, across the entire solution time (adults: $M = 0.34$, $SD = 0.05$; children: $M = 0.33$, $SD = 0.06$), and there was no significant difference between groups ($\beta = -0.01$, $SE = 0.01$, $t = -0.75$, $p = .46$). Focusing just on initial solution time (the first third of the trial) revealed that a higher proportion of looking to these items occurred when participants first viewed these problems than across the entire trial with adults and children spending nearly half of initial problem-solving time focused on this relation (adults: $M = 0.53$, $SD = 0.09$; children: $M = 0.52$, $SD = 0.11$). Again, there was no significant difference in this looking pattern between groups ($\beta = 0.02$, SE $= 0.02$, $t = -0.85$, $p = .40$).

In order to make conclusions about whether looking to the source relation supports successful reasoning, we assessed the relation between performance and visual attention patterns. Unsurprisingly, children performed poorly on scene analogies, answering less than one-third of the problems correctly ($M = 0.30$, $SD = 0.26$), whereas adults were much more accurate ($M = 0.92$, $SD = 0.06$). Because adults performed nearly at ceiling, we only assessed whether looking patterns predict accuracy for child participants.

Binomial generalized linear models, with accuracy on each problem (0,1) as the dependent measure, were used to determine whether looking to source relation is predictive of behavioral performance. In contrast to prior work, we found no relation between performance and looking to the source relation across the entire trial ($\beta = 0.07$, $SE = 0.81$, $t = 0.08$, $p = .93$), or during initial problem solving ($\beta = 0.41$, $SE = 0.46$, $t = 0.89$, $p = .37$) and learning. Overall, these results challenge previous work suggesting that children's lower performance on analogy problems can be explained by failures to attend adequately to the source relationship.

3.2. Featurally salient distractors

When children and adults were not looking at the source relation, how did they allocate their attention? In line with previous work assessing children's visual attention during analogical reasoning, children spent roughly 10% percent more of their time looking to featural distractors ($M = 0.20$, $SD = 0.10$) as compared to adults ($M = 0.09$, $SD = 0.02$) across the entire problem-solving time, and this difference was statistically significant ($\beta = 0.09$, $SE = 0.01$, $t = 6.55$, $p < .001$).[7] This difference was also observed during initial problem solving (adults: $M = 0.08$, $SD = 0.04$; children: $M = 0.11$; $SD = 0.05$; $\beta = 0.03$, $SE = 0.01$, $t = 3.05$, $p = .003$). This indicates that children allocate a higher proportion of their visual attention to the distractor than adults initially and throughout the course of the solving process.

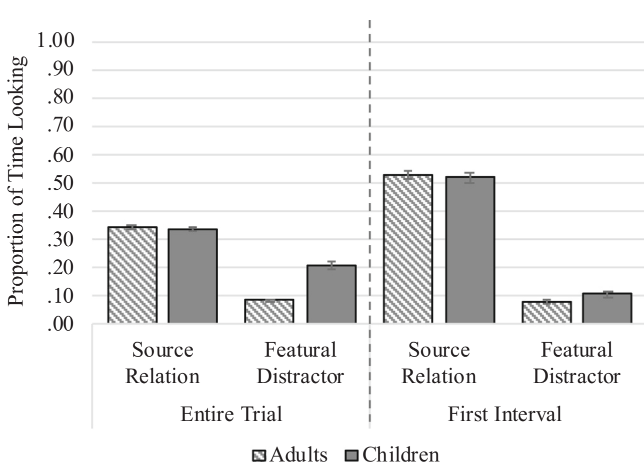

Fig. 3. Proportion of time looking to source relations and distractors across entire trials and the first interval of trials.

Finally, we asked whether behavioral accuracy was predicted by looking to the distractor. Interestingly, in line with previous work, children performed better if they spent less time looking to the distractor across the entire problem-solving time ($\beta = -9.80$, $SE = 1.23$, $t = -7.61$, $p < .001$). However, when considering initial looking times only, there was no relation between looking to the distractor and performance ($\beta = -1.22$, $SE = 0.78$, $t = -1.57$, $p = .12$), suggesting that initial attention to the scenes was not the key differentiating period predicting success, but that if a child could not disengage his or her attention from the distractor across the problem-solving time, this reliably led to them incorrectly selecting the featural distractor.

3.3. Strategy classification

To further compare visual attention patterns employed by children and adults during analogical reasoning, we adapted the classification procedure developed by Starr et al. (2018) and Vendetti et al. (2017). Participants' looking patterns were classified as following one of three problem-solving strategies proposed in the literature (i.e., projection-first, structure-mapping, semantic-constraint; see Table 1). In line with previous work using propositional analogies, just under half of adult trials were classified as projection-first ($M = 0.44$, $SD = 0.16$), one third of trials were classified as structure-mapping ($M = 0.34$, $SD = 0.13$), and the remainder of trials were classified as semantic-constraint ($M = 0.22$, $SD = 0.11$; Fig. 4). This suggests that adults employ similar strategies across analogy formats. However, we find notable differences between the strategies employed by children when solving scene analogies (current study) versus propositional analogies (Starr et al., 2018). Starr and colleagues found that that children were more likely to use a semantic-constraint than projection-first strategy. For scene analogies, we found that children employed the projection-first ($M = 0.35$, $SD = 0.20$) and semantic-constraint ($M = 0.40$, $SD = 0.18$) strategies on a similar number of trials. We also found that children utilized a structure-mapping strategy for ~25% of their trials ($M = 0.25$, $SD = 0.16$), but cannot consider how this differs from structure-mapping strategy use for children solving propositional analogies, as Starr et al. (2018) did not consider this strategy.

To determine whether adults and children show significant differences in strategy use, we used a general linear model with proportion of trials on which a strategy was employed as the dependent variable, age group (child, adult), strategy (projection-first, semantic-constraint, structure-mapping), and an interaction between age group and strategy as predictors, and a random effect of participant. Results are interpreted within the context of a significant interaction between age group and strategy ($\chi^2(2) = 14.17$, $p < .001$). Follow-up analyses showed that

[7] On average, both children and adults look to the featural distractor slightly more on incorrect problems (children: $M = 0.23$, adults: $M = 0.14$) rather correct problems (children: $M = 0.09$, adult: $M = 0.08$). However, children and adults look to the source relation similarly despite solving problems incorrectly (children: $M = 0.34$, adults: $M = 0.36$) or correctly (children: $M = 0.34$, adults: $M = 0.35$). Similarly during the first interval, children and adults look to the featural distractor slightly more on incorrect problems (children: $M = 0.12$, adults: $M = 0.10$) rather correct problems (children: $M = 0.08$, adult: $M = 0.07$). Children and adults also look to the source relation similarly despite solving problems incorrectly (children: $M = 0.50$, adults: $M = 0.49$) or correctly (children: $M = 0.52$, adults: $M = 0.53$).

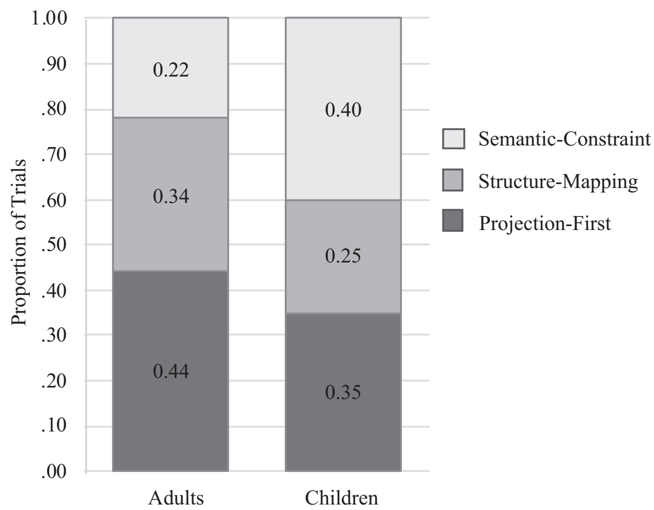

Fig. 4. Proportion of trials classified as each strategy for adults versus children.

whereas children use semantic-constraint strategies significantly more than adults ($\beta = 0.11$, $SE = 0.03$, $t = 3.82$, $p < .001$), there were no difference between children and adults in use of projection first ($\beta = -0.06$, $SE = 0.04$, $t = -1.41$, $p = .16$) or structure-mapping ($\beta = -0.01$, $SE = 0.03$, $t = -0.52$, $p = .60$) strategies.

Next, we assessed how strategy classification was related to performance on a trial-by-trial basis. Because adults were nearly at ceiling in terms of performance, we only consider how child classification was related to performance (Fig. 5). Based on prior work using propositional analogies (Starr et al., 2018), we anticipated that use of a projection-first strategy would lead to successful problem-solving. Using binomial logistic regression models with trial performance (0,1; accuracy or distractor error) as the dependent variable and trial classification as the predictor, with participant as a random effect, models revealed no significant effects of trial classification on accuracy ($\chi^2(2) = 2.30$, $p = .32$) or choice of distractor ($\chi^2(2) = 0.38$, $p = .83$). In contrast to prior work, this suggests that children's pattern of looking during solving does not predict their ability to avoid the featural match and arrive at the correct, relational match.

One reason our findings do not align with previous work may be that the classification equations do not represent the three strategies in the same way across formats. To address this, we considered the visual attention patterns that would *hypothetically* occur in trials classified as each strategy, as proposed by the theoretical models driving each strategy: In projection-first and structure-mapping strategies we would *expect* more time looking to the items involved in the relations, and in semantic-constraint strategies we would *expect* more time looking to response options, in particular the featural match. Using a pair of general linear models, we asked if strategy classification was related to proportion looking to two AOIs: the source relation and the featural distractor. Models with proportion of time looking to AOI (source relation or distractor) as the dependent measure and trial classification as the predictor, with participant as a random effect, revealed that trial classification significantly related to time looking to the distractor ($\chi^2(2) = 11.46$, $p = .003$) and to the source relation ($\chi^2(2) = 111.67$, $p < .001$). Follow-up contrasts indicate strategy classification was related to proportion of time looking to individual AOIs. Children looked *more* to the distractor when employing a semantic-constraint strategy than a projection-first ($\beta = 0.01$, $SE = 0.01$, $t = 2.67$, $p = .008$) or structure-mapping strategy ($\beta = 0.04$, $SE = 0.01$, $t = 2.85$, $p = .005$), and there was no difference in overall looking time when using a projection-first or semantic-constraint strategy ($\beta = -0.00$, $SE = 0.01$, $t = -0.47$, $p = .64$). Children looked *less* to the source relation when employing a semantic-constraint strategy than a projection-first ($\beta = -0.11$, $SE = 0.01$, $t = -10.40$, $p < .001$) or structure-mapping strategy ($\beta = 0.00$, $SE = 0.00$, $t = 7.23$, $p < .001$), and *more* when using a projection-first than structure-mapping strategy ($\beta = -0.03$, $SE = 0.01$, $t = -2.12$, $p = .04$). These findings replicate prior work (Starr et al., 2018) demonstrating that the time children spend looking to particular components of analogy problems relates to strategy use, even though in the present study, strategy is not predictive of performance.

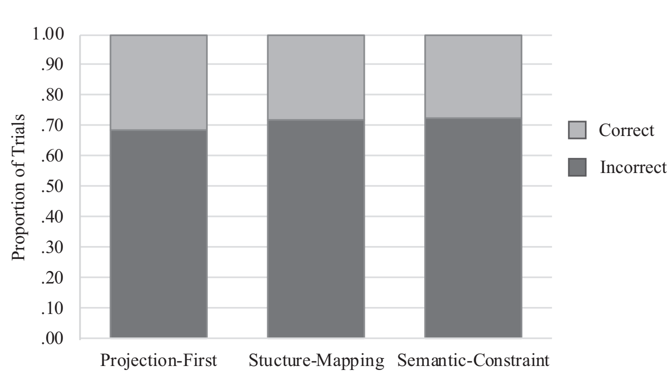

Fig. 5. Proportion of trials correct versus incorrect split by trial classification.

Therefore, we suggest that while this classification procedure is suitable for considering visual attention patterns during scene analogies, in that the equations represent the three theoretical models of analogical reasoning, these strategies do not necessarily reflect the answer a child will reach while problem solving. Instead, the amount of looking to the featural distractor was the only measure in our study that reliably predicted children's behavioral performance.

4. Discussion

Whereas previous research using propositional analogies suggests the mechanism underpinning children's primary errors during analogical reasoning is a failure to attend to relational information, our results show that this conclusion is *not* true of all analogy-problem formats. When children and adults solve scene analogies, attention to the source relation during initial problem solving and across the full problem-solving period does not differ between age groups, nor does attention to this relation predict accuracy in children. In contrast, we show that attention to featural distractors *does* affect problem-solving: Children look more to the distractor, and attention to the featural distractor predicted accuracy in children. And, considering our results in relation to prior work (e.g., Starr et al., 2018; Thibaut & French, 2016), we find that whereas the visual strategies adults use during solving do not differ across analogy format, the organization of children's attention is impacted by format. During a propositional analogy task, children were most likely to use a semantic-constraint strategy when solving problems, and this predicted accuracy (Starr et al., 2018). During our scene analogy task, children were just as likely to use a semantic-constraint strategy or a projection-first strategy, and strategy *did not* predict accuracy. Together, these results suggest the format of the problem has a serious impact on children's visual attention patterns during analogical reasoning, and that prioritization of relational information is not always critical for successful problem solving across all analogy paradigms. We discuss *why* the format of propositional versus scene analogy problems might lead to these different findings and the practical implications of this work for educators.

Our findings that analogy format impacts children's performance and visual attention patterns are in line with previous work suggesting that children's ability to problem solve is influenced by the structure of a given problem (e.g., Cantlon et al., 2007; Mix, 1999; Paik & Mix, 2006). While both propositional and scene analogies require processing relational information in order to arrive at a correct solution, their structures differ significantly, which may account for the differences in how children approach these problems. In analogies of A:B::C:D format, children who are not skilled at analogical reasoning seem to overlook the relational information contained in the A:B pair, and focus on the C item and response options. They interpret the task as 'match C to something' and treat A:B as irrelevant. In the example used previously, children might ignore that *likes to eat* is the relation shared between the elephant (item B) and the peanut (item A), and be more likely to pick an option that is similar to the banana (item C) on another dimension, such as *color* or *shape*. It seems that the salience of the A:B pair is not great enough to warrant attention from those children who do not understand the task goal.

In contrast, in a scene analogy, children's visual attention is drawn to the source relation initially, whether or not they understand its importance. This is because the circled item is part of the source relation. Based on our results, the salience of the circled item draws both children and adults' attention equally at first, but unlike adults, children less often utilized that information to correctly solve the problem – just because children visually attend to the items does not mean they process the deeper, relational context necessary for successful analogical reasoning.

This problem may be endemic to the format of the propositional analogies. Propositional analogies, even when solved relationally, can be solved as a series of two relational discovery problems: 1) discover the relation between A:B, and 2) find that relation between C and an answer choice. Reasoners are not required to align A and C, or hold an A:C relation in mind while reasoning about C:D – rather they can simply use the C and the relation to find the solution option. Thus, if scene analogies require more relational reasoning, it is reasonable to expect that one would reveal the more potent role of featural distractors on analogical thinking in that setting. While both problem formats present their challenges for young solvers, it seems that the presence of distractors is the dominant struggle when they are solving scene analogies. This is important information for educators and researchers, who may consider incorporating problems of analogy in their classroom instruction or experiments, respectively. Being aware of the impact of problem format could answer questions about variability in children's performance across tasks, or even across domains.

Whereas visual attention to the source relation did not predict accuracy, attention to the featural distractor was found to relate to performance. Across the entire problem-solving period children allocated more of their attention to the distractor than adults, and this was negatively related to behavioral performance. Interestingly, while children also *initially* look more to the distractor, this was *not* predictive of performance. Corroborating behavioral work (Richland et al., 2006), this suggests that it is the inability to disengage from featural distractors across the problem-solving period that leads to poor performance – children who get visually 'stuck' on the distractor are more likely to incorrectly solve the problem. Although Starr and colleagues find that prioritization of the source relation is the most important predictor of correct problem-solving in a propositional analogy task, our findings do not stand in complete opposition to this work: Starr and colleagues still show that children attend more to the featural distractor than adults, it is just not the most important factor in determining successful problem-solving. This suggests that across analogy formats, although the featural distractor may be a common pitfall, surface-level features of the problem influence whether attention to the source relation matters. In contrast to the differences seen in how children visually attend to analogies of different formats and the effect of looking patterns on accuracy, we found that adults perform very similarly on scene analogy problems in the current study as they do on propositional analogies in previous work (Vendetti et al., 2017), spending roughly a third of their time looking towards the source relation or the A and B items. Perhaps because adults understand that they need to identify the deeper structure in these problems, and therefore, are more proficient analogical reasoners, they are not restricted by the structure of the problem.

As a secondary aim of this work, we investigated whether strategy use observed for propositional analogies (Starr et al., 2018; Vendetti et al., 2017) would be conserved for scene analogies, and would predict performance for children as it did for propositional analogies. While adults' visual attention patterns seem consistent across analogy formats, children's attentional patterns seem to differ between propositional and scene analogy formats. In contrast to prior work with propositional analogies showing that children are most likely to use a semantic-constraint strategy and this negatively relates to performance (Starr et al., 2018), here, we found that children are just as likely to use a semantic-constraint or projection-first strategy and that strategy use does not relate to children's ability to solve scene analogies. Again, these results speak to the importance of considering the structure of a problem: In a propositional analogy, there is not a particularly salient aspect of the problem structure that would initially draw children's attention to the A:B pair. Whereas, in a scene analogy, the circled, prompted item *is in* the source relation. Given that previous work consistently demonstrates that children are drawn to featural matches (e.g., Richland et al., 2006; Thibaut & French, 2016), it is understandable that children's attention may be differentially impacted by problem format in the pursuit of the featural match. That is, when children focus on the immediate task goal in a propositional analogy (i.e., 'Find something that goes with C') the featural match is between the C item and a response option. Therefore, there is no need to revert attention to the A:B pair. However, in a scene analogy, in order to make a featural match between an item in the target scene and the circled item in the source scene, children are required to attend to the source relation, even if just to the prompted item. Thus, it is perhaps unsurprising that children would look to the source relation first and be classified as using a projection-first strategy, which would account for children's lack of difference between semantic-constraint and projection-first strategy use. But importantly, even if children look to the source relation, and are classified as using a projection-first strategy, this does not predict arriving at the correct solution. Children can be shown where to look, but not understand the *relevance* of the source relation and not use it to correctly solve a scene analogy.

Although our work adds an important piece to understanding the development of analogical reasoning ability, two limitations should be noted. First, one limitation of this work lies in the restricted comparisons that can be made between propositional and scene analogies. Specifically, previous analogical reasoning research has made strong conclusions about the differences between age groups in terms of looking to the C item in propositional analogies (A:B::C:D), such that children look more to the C item earlier in the problem solving process and focus their search around C, whereas adults will search in a more organized way by first examining the A:B pair and then looking at the C item and the possible answers (Starr et al., 2018; Thibaut & French, 2016). Unfortunately, there is not a functionally comparable item to C in a scene analogy. Items C and D are already in relation with one another in a scene analogy, whereas D must be chosen from multiple options by the participant in a propositional analogy. Therefore, in this study, it is difficult to make conclusions about looking to the C item. And while this is true, using the classification procedure, we did choose to consider the prompted, circled item the functional equivalent to the C item. Both items have an associated featural match within the problem, and can, thus, be treated as comparable. However, we acknowledge that this is not a perfect comparison. This, again, speaks to the structural difference between propositional analogies and scene analogies and the limitations in making comparisons across problem formats. Another way that we aim to expand this work in the future is to examine additional relations and problem contexts to better understand the generality of children's attentional patterns and relational thinking.

A second limitation lies in our ability to make comparisons to previous work with propositional analogies due to differences in analytical approaches. In the present study, we considered how looking patterns predicted a child's ability to solve scene analogy problems. In order to understand the differences between looking patterns that lead children to *correct* choices versus *incorrect* choices, we included trials on which children answered correctly and incorrectly. Although this provides important variability for predicting which looking patterns lead children to make an incorrect selection, this approach departs from that of previous work. Importantly, Starr and colleagues focused their analysis on children's correct trials, rather than including incorrect trials. While comparisons between Starr et al. and the present study are necessary to consider the potential impact of problem structure, a degree of caution is warranted. Because of this difference between the present study and Starr et al.'s work it is unclear whether differences in children's strategy use directly reflects differences due to problem format or differences in the strategies used for correct vs. incorrect choices. Future work should explore these uncertainties more closely, either by isolating analogical reasoning problems solved correctly to explore whether children's strategy use differs due to problem format or explore the relation between accuracy (and errors) and strategy use during solving of propositional analogies. Either of these approaches would build from the present study's findings to address the root cause of differences in strategy use.

Despite some limitations, overall, this work contributes to the growing body of literature employing eye tracking to elucidate the attentional mechanisms underlying learning in analogical reasoning. Results suggest that although there are some differences between how adults and children process analogies that are generalizable across formats, there are other aspects of how attention is allocated that are dependent upon analogy type – particularly with young children. Problem format matters and influences the type of errors children make – a conclusion that has relevance both for researchers using analogical reasoning tasks in the lab, and educators who may teach analogical reasoning through a number of different formats. Specifically, it is important for educators to understand the impact that problem structure or task design has on young children in learning contexts and be aware that children's difficulties during tasks of relational reasoning may vary based on problem or task type. While this study does not directly address the developmental change that contributes to improvements in reasoning, we suggest that as children develop they are increasingly able to apply their visual attention in appropriate ways to use analogy for understanding relational information. Given that analogical reasoning plays a central role in human thought in a variety of contexts, developing techniques to support this ability in young children is essential to support their future success.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

There are no known conflicts of interest related to this manuscript.

References

Brand, C. O., Mesoudi, A., & Smaldino, P. E. (2021). Analogy as a catalyst for cumulative cultural evolution. *Trends in Cognitive Science, 21*, S1364–S6613. https://doi.org/10.1016/j.tics.2021.03.002

Cantlon, J., Fink, R., Safford, K., & Brannon, E. M. (2007). Heterogeneity impairs numerical matching but not numerical ordering in preschool children. *Developmental Science, 10*, 431–440. https://doi.org/10.1111/j.1467-7687.2007.00597.x

Chalmers, D. J., French, R. M., & Hofstadter, D. R. (1992). High-level perception, representation, and analogy: A critique of artificial intelligence methodology. *Journal of Experimental & Theoretical Artificial Intelligence, 4*, 185–211. https://doi.org/10.1080/09528139208953747

Chen, Z., Honomichl, R., Kennedy, D., & Tan, E. (2016). Aiming to complete the matrix: eye-movement analysis of processing strategies in children's relational thinking. *Developmental Psychology, 52*, 867–878. https://doi.org/10.1037/dev0000113

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*, 1–43. https://doi.org/10.1037/0033-295X.115.1.1

Doumas, L. A. A., Morrison, R. G., & Richland, L. E. (2018). Individual differences in relational learning and analogical reasoning: A computational model of longitudinal change. *Frontiers in Psychology, 9*, 1–14. https://doi.org/10.3389/fpsyg.2018.01235

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*, 1–63. https://doi.org/10.1016/0004-3702(89)90077-5

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141–205. https://doi.org/10.1207/s15516709cog1902_1

French, R. M., & Thibaut, J. P. (2014). Using eye-tracking to predict children's success or failure on analogy tasks. In *Proceedings of the Thirty-Sixth Annual Meeting of the Cognitive Science Society* (pp. 2222–2227).

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170. https://doi.org/10.1207/s15516709cog0702_3

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47–59. https://doi.org/10.2307/1130388

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*, 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gentner, D., Holyoak, K. J., & Kokinov, B. N. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.

Glady, Y., French, R. M., & Thibaut, J. P. (2016). Comparing competing views of analogy making using eye-tracking technology. In *Proceedings of the thirty-eighth annual meeting of the Cognitive Science Society* (pp. 1349–1354).

Glady, Y., French, R. M., & Thibaut, J. P. (2017). Children's failure in analogical reasoning tasks: A problem of focus of attention and information integration? *Frontiers in Psychology, 8*, 1–13. https://doi.org/10.3389/fpsyg.2017.00707

Glady, Y., Thibaut, J. P., & French, R. (2010). Visual strategies in analogical reasoning development: A new method for classifying scanpaths. *Proceedings of Thirty-Fifth Annual Meeting of the Cognitive Science Society*, 2398–2403. https://doi.org/10.13140/2.1.4107.1365

Gordon, P. C., & Moser, S. (2007). Insight into analogies: Evidence from eye movements. *Visual Cognition, 15*, 20–35. https://doi.org/10.1080/13506280600871891

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 234–259). New York: Oxford University Press.

Holyoak, K. J., Junn, E. N., & Billman, D. (1984). Development of analogical problem-solving skill. *Child Development, 55*, 2042–2055. https://doi.org/10.2307/1129778

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Hosenfeld, B., van der Maas, H. L. J., & van den Boom, D. (1997). Indicators of discontinuous change in the development of analogical reasoning. *Journal of Experimental Child Psychology, 64*, 367–395.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466. https://doi.org/10.1037/0033-295X.104.3.427

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220–264. https://doi.org/10.1037/0033-295X.110.2.220

Leung, S. S., & Silver, E. A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers. *Mathematics Education Research Journal, 9*, 5–24. https://doi.org/10.1007/BF03217299

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32*, 517–535. https://doi.org/10.1006/jmla.1993.1027

Markman, A. B., & Wood, K. (2009). *Tools for innovation: The science behind the practical methods that drive new ideas*. New York, NY: Oxford University Press.

Mix, K. S. (1999). Similarity and numerical equivalence. *Cognitive Development, 14*, 269–297. https://doi.org/10.1016/s0885-2014(99)00005-2

Morrison, R. G., Doumas, L. A. 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. https://doi.org/10.1111/j.1467-7687.2010.00999.x

Paik, J. H., & Mix, K. S. (2006). Preschoolers' use of surface similarity in object comparisons: Taking context into account. *Journal of Experimental Child Psychology, 95*, 194–214. https://doi.org/10.1016/j.jecp.2006.06.002

Rattermann, 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. https://doi.org/10.1016/S0885-2014(98)90003-X

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. https://doi.org/10.1016/j.jecp.2006.02.002

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 177–192. https://doi.org/10.1002/wcs.1336

Siegel, L. S. (1974). Heterogeneity and spatial factors as determinants of numeration ability. *Child Development, 45*, 532–534. https://doi.org/10.2307/1127984

Snow, R. E., Kyllonen, C. P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 47–103). Hillsdale, NJ: Erlbaum.

Starr, A., Vendetti, M. S., & Bunge, S. A. (2018). Eye movements provide insight into individual differences in children's analogical reasoning strategies. *Acta Psychologica, 186*, 18–26. https://doi.org/10.1016/j.actpsy.2018.04.002

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

Thibaut, J. P., French, R., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive functions. *Journal of Experimental Child Psychology, 106*, 1–19. https://doi.org/10.1016/j.jecp.2010.01.001

Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development, 38*, 10–26. https://doi.org/10.1016/j.cogdev.2015.12.002

Vendetti, M. S., Starr, A., Johnson, E. L., Modavi, K., & Bunge, S. A. (2017). Eye movements reveal optimal strategies for analogical reasoning. *Frontiers in Psychology, 8*, 2–10. https://doi.org/10.3389/fpsyg.2017.00932

Walley, A. C., Michela, V. L., & Wood, D. R. (1995). The gating paradigm: Effects of presentation format on spoken word recognition by children and adults. *Perception & Psychophysics, 57*, 343–351. https://doi.org/10.3758/BF03213059