



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

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

Exploring how visual attention, inhibitory control, and co-speech gesture instruction contribute to children's analogical reasoning ability

Katharine F. Guarino^{a,*}, Elizabeth M. Wakefield^a, Robert G. Morrison^a,
Lindsey E. Richland^b

^a Loyola University Chicago, United States

^b University of California, Irvine, United States

ARTICLE INFO

Keywords:

Analogy
Attention
Eye tracking
Inhibitory control
Gesture
Reasoning

ABSTRACT

Analogical reasoning is an important, but difficult skill to develop. Here, we consider (1) whether children's difficulty with analogical reasoning can be understood by measuring visual attention during analogy problem solving and a measure of inhibitory control, which researchers have suggested plays an integral role in analogical reasoning, and (2) whether analogical reasoning can be supported with co-speech gesture instruction. Children (4–5 year-olds) solved scene analogies before and after speech-alone or speech + gesture instruction. Results suggest relations exist between some behavioral performance measures, inhibitory control, and visual attention, but some of these relations change after instruction. During instruction, gesture led to more mature looking patterns, although this did not predict performance: All children showed gains from pretest to posttest. We consider how these findings extend prior gesture and analogical reasoning literature, call for replication of these findings, and identify future directions of research.

1. Introduction

Analogical reasoning – the ability to identify underlying schematic structure shared between experiences – is one of the most important ways children make sense of the world around them (e.g., Ferry, Hespos, & Gentner, 2015; Goswami, 2001; Rattermann & Gentner, 1998). Research shows that mature analogical reasoning contributes to human innovation (Markman & Wood, 2009), creativity (Holyoak & Thagard, 1995; Sternberg, 1988), inductive and scientific reasoning (Dunbar & Klahr, 2012; Gentner, 2010), and children's ability to learn within a variety of instructional domains, including mathematics, science, and history (Richland & Simms, 2015); however, this skill is difficult to acquire, showing a protracted development across early childhood (Gentner & Smith, 2013).

Analogical reasoning requires an individual to look beyond *featural* similarities between elements in two domains and notice the shared *relational* structures between the domains. But, even when told explicitly to look for relational similarities to solve problems, young children tend to focus on featural similarities. This has been demonstrated in lab settings using a scene analogy task developed by Richland et al. (2006). In this task, children examine two scenes (i.e., a source and target scene) which contain both relational (abstract) similarities and featural similarities. Children are asked to identify an item in the target scene that corresponds *relationally* to a prompted item in the source scene, but often choose an item that corresponds *featurally* to the prompted item instead. For example, a

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

<https://doi.org/10.1016/j.cogdev.2021.101040>

Received 4 June 2020; Received in revised form 4 February 2021; Accepted 9 March 2021

Available online 21 March 2021

0885-2014/© 2021 Elsevier Inc. All rights reserved.

source scene might show a dog chasing a cat, with the dog prompted. The corresponding target scene would contain a boy chasing a girl and a second dog. Here, the boy would be the correct relational choice, and the dog would be the incorrect featural distractor. Whereas adults and older children can identify the correct relational choice despite the distractor – choosing the boy in the above example, 3- and 4-year-olds are much more likely to choose the featural distractor (Richland, Morrison, & Holyoak, 2006). If the featural distractor is removed from the target scene young children are able to choose the relational choice.

Part of young children's difficulty with analogical reasoning may stem from insufficient inhibitory control (IC), a type of executive function (e.g., Dumas, Morrison, & Richland, 2018; Simms, Frausel, & Richland, 2018). Whereas the role of IC in adult analogical reasoning has long been established (e.g., Morrison et al., 2004; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004), there is a growing body of work that posits that IC is also important for the development of analogical reasoning in children (e.g., Begolli et al., 2018; Dumas et al., 2018; Morrison, Dumas, & Richland, 2011; Richland, Chan, Morrison, & Au, 2010, 2006; Thibaut, French, & Vezneva, 2010). According to Morrison et al. (2011), adult's IC may act as a gating mechanism for an individual's working memory. When faced with a scene analogy like the one described above, individuals activate featural and relational knowledge stored in long-term memory to make sense of the scenes and complete a task. Considering the example above, this activated information will allow them to choose an item in the target scene that is in the same place in the pattern as the dog in the source scene. IC determines what information will remain active in working memory, and therefore, helps to determine which item an individual will choose. By this theory, adults, who have higher IC, are able to more systematically gate relational knowledge in working memory which helps them complete the task – in this case, understanding the idea of chasing, and holding that in working memory while determining that the most relevant relational correspondence is that the dog in the source scene and the boy in the target scene are both chasing.

In contrast, if an individual has low IC, the gating mechanism does not function as efficiently, and featural information that is most saliently related to the prompted item, the dog, will remain activated and dominate working memory. This theory would suggest this is the case for young children, who would then be less able to inhibit activation of the concept 'dog' and therefore choose the salient, featural distractor. Age has been shown regularly to predict the likelihood of selecting a distractor in scene analogy tasks, corroborating this account (Murphy, Zheng, Shivaram, Vollman, & Richland, 2021; Richland et al., 2006, 2010). Further, the argument that IC, and executive functions (EFs) more broadly, are important predictors of analogical reasoning ability has been corroborated across a number of methodological approaches including computational modeling (Dumas et al., 2018; Morrison et al., 2004, 2011; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004), longitudinal work (Richland & Burchinal, 2013; Richland & Simms, 2015), and eye tracking (e.g., Glady, Thibaut, & French, 2010; Starr, Vendetti, & Bunge, 2018). This line of works suggests that over development, as EFs mature, children are increasingly able to select relational matches and avoid selecting featural matches.

Whereas previous work has implicated IC as an important contributor to mature analogical reasoning ability, there are still many unknowns about how IC relates to other behavioral measures which are relevant for problem-solving, and whether children can benefit from directed instruction despite immature IC. In the present study, we consider: (1) whether IC might help children correctly solve analogical reasoning problems by promoting visual attention away from featural distractors, and (2) whether analogical reasoning ability could be facilitated through instruction using gesture, a tool that has been shown to promote learning in a variety of domains (e.g., Congdon, Kwon, & Levine, 2018; Cook, Mitchell, & Goldin-Meadow, 2008; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014; Ping & Goldin-Meadow, 2008; Valenzeno, Alibali, & Klatzky, 2003; Wakefield, Hall, James, & Goldin-Meadow, 2018; Wakefield & James, 2015).

1.1. The role of visual attention in analogical reasoning

In previous work, researchers have established that children and adults who are successful when solving analogical reasoning problems tend to show different patterns of visual attention than children who unsuccessfully choose the featural distractor. For example, propositional analogies in the format of A:B::C:D require individuals to identify the relation between A and B, and then choose 'D' from an array of options, such that C and D relate to each other in the same way that A and B relate to each other. When successfully solving these analogies, children and adults tend to initially direct their visual attention to the A:B pair before attending to the C item and response options (Starr et al., 2018; Thibaut & French, 2016). When unsuccessfully solving A:B::C:D analogies, children tend to show a lack of attention to relations within analogies and a constrained focus on the C item – scanning between the C item and response options.

When considering visual attention to scene analogy problems, Guarino et al. (2019) also found differences in how adults and children allocate their attention: 4–5 year-old children attend more to the featural match than adults. Taken together, we see different visual attention patterns in adults who successfully solve problems of analogy versus children who fail to correctly solve problems of

analogy, although the specific differences in attention are affected by problem format. While orienting attention away from featural matches is important for successfully solving either form of analogy, attention to the source relation depends on the format of the problem: Attention to the source relation in a propositional analogy (A:B pair) predicts children's performance (Starr et al., 2018), yet attention to the source relation in a scene analogy (items involved in the relation of interest in the source scene) does not predict performance (Guarino, Wakefield, Morrison, & Richland, 2019).

What might the differences in visual attention to propositional and scene analogy problems mean? We suggest that in each case, the patterns of visual attention align with what we might expect from children with higher versus lower IC, and thus, looking patterns may be driven by this important predictor of analogical reasoning ability. In the case of propositional analogies, children with higher IC theoretically would be able to inhibit attention away from the immediate task goal (i.e., 'Find what goes with C') allowing them to first assess the relation between the A:B pair before reviewing response options. In contrast, children with lower IC who do not inhibit that task goal might instead jump to the C item and make choices based on featural similarities between response options and the C item (i.e., choosing a featural distractor because it is featurally similar to the C item), rather than consider the meaningful relation between the A:B pair. Indeed, children who have difficulty inhibiting the immediate task goal focus more of their visual attention on the featural distractor compared to other response options (Starr et al., 2018). Similarly, when solving a scene analogy, children with lower IC will likely be unable to inhibit attention away from the item in the target scene that is featurally similar to the prompted item, in favor of focusing on relationally similar items in the problem.

In the present study, we will *directly* compare the relation between IC and visual attention during analogical reasoning to test this relationship using scene analogy problems. To measure IC, we will use a task appropriate for children, that has been used in previous research on analogical reasoning ability (Simms et al., 2018), the Eriksen flanker task, based on the Attention Network Task (Rueda et al., 2004). In analogical reasoning tasks, children must ignore featural similarities in favor of relational similarities. In other words, they must focus their attention on task-relevant features, and inhibit other features. The flanker task also requires children to focus their attention on task-relevant features, and inhibit other features in order to make the appropriate behavioral response. Specifically, in this version of the task they must determine the direction a center fish is 'swimming' (i.e., facing), embedded within a line of fishes. Whereas congruent trials show all fish swimming in the same direction, on incongruent trials, the fish surrounding the center fish are swimming in the opposite direction of the center fish. Thus, children must ignore the more salient information of the direction most fish are swimming, in favor of the center fish. While recent work suggests that children's performance on the flanker task does not relate to their ability to solve analogical reasoning problems, Simms et al. (2018) suggested that this might have been because a wide age-range had been considered. The researchers suggest that at a younger age – the age we use in the current study – this task may be a good predictor of analogical reasoning ability.

1.2. Gesture instruction may be one way to support analogical reasoning ability

In addition to considering the relation between a child's IC and visual attention during scene analogy problem solving, we ask whether co-speech *gestures*, movements of the hands that are produced simultaneously with speech, represent information (McNeill, 1992), and direct visual attention (e.g., Rohlfing, Longo, & Bertenthal, 2012; Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018), can serve as a tool to support children's analogical reasoning through facilitating effective looking patterns (i.e., directing visual attention away from featural distractors and towards important, relational comparisons). Decades of research suggest that gesture helps children learn new concepts in a variety of domains, from mathematical principles (e.g., Cook, Duffy, & Fenn, 2013) to word learning (e.g., Wakefield, Hall et al., 2018). Although the mechanisms by which gesture facilitates learning are still being understood, there is evidence that gesture can organize children's visual attention (Wakefield, Novack et al., 2018): Gesture use by an instructor increases the likelihood that children will focus on the important aspects of a problem or material being taught, and better align their visual attention with referents in a teacher's spoken instruction, which is predictive of learning gains (Wakefield, Novack et al., 2018). In the case of an analogical reasoning problem, gesture could be used to call attention to and link related items as they are mentioned in spoken instruction. The benefits of this could be two-fold: First, promoting comparison between exemplars by visually aligning them through gesture could encourage children to attend to the relational or structural similarities shared among them (Namy & Gentner, 2002). Second, by drawing attention to the abstract relations in the source and target scene, this necessarily draws attention *away* from the featural distractor.

1.3. Present study

In the present study, we expand on previous literature by considering how children's visual attention when solving scene analogy problems may be related to their IC, and whether gesture may be a useful tool for supporting successful problem solving, potentially through its ability to direct visual attention. To address these questions, we use a pretest-training-posttest design and monitor children's visual attention using eye tracking while they solve problems and receive instruction through speech alone or speech and gesture, and measure their IC using the Eriksen flanker task.

We hypothesize that individual differences in IC, as measured by the Eriksen flanker task, will be related to both behavioral performance and visual attention patterns as measured using eye tracking before children have received instruction. By measuring visual attention via eye tracking we can assess how children orient their foveal attention while solving scene analogies. Specifically, at pretest, the higher a child's IC, the more likely they should be to choose the correct, relational choice and the less likely they should be to choose the featural distractor. Further, the higher a child's IC, the less likely they should be to visually attend to the featural distractor. Based on previous literature from other domains, we anticipate that children will benefit more from instruction that includes gesture, than from instruction through speech alone, and that this may be driven by visual attention patterns. Gesture instruction may scaffold understanding of the relations through helping children avoid the featural matches and direct attention towards relational information in the problem.

2. Method

2.1. Participants

Fifty-seven 4- and 5-year-old children participated in the present study (29 females, $M_{\text{age}} = 4;11$ mo, $SD_{\text{age}} = 5.6$ mo)¹. Children were randomly assigned to one of two instruction conditions: speech-alone ($n = 28$) or speech + gesture ($n = 29$). Two children in the speech + gesture condition were excluded from all analyses due to a lack of behavioral response. Participants represented a diverse sample from a large metropolitan city (48 % White, 14 % Black, 5% More than one race, 5% Asian, 28 % Unreported). Informed consent was obtained from a parent or guardian of each participant, and verbal assent was obtained from children. Children were compensated with stickers and a certificate noting their participation in a research study. Additionally, study sites were compensated with a gift card to purchase classroom materials. Children participated individually in one experimental session at their school during a regular school day.

2.2. Materials

2.2.1. Pretest/posttest stimuli

Twenty-four scene analogy problems were created, based on the structure used by Richland and colleagues (2006). Each problem included a pair of scenes, a source scene on the left, and a target scene on the right. Scenes depicted one of two *relation categories* (i.e., chasing or reading) occurring between items (i.e., animals or people). Source scenes contained five items: the two items within the relation of chasing or reading, and three additional items (i.e., neutral inanimate objects that were not involved in the relation of interest). One of the items within the source scene relation was circled. Target scenes also contained five items: the two items within the relation, two additional items, and a featural distractor. The featural distractor was similar to the circled source-scene item and centrally located, increasing the likelihood that the item would draw participants' attention.

Fig. 1 A shows an example of a chasing *source* and *target* scene. The source scene on the left shows a tiger chasing a woman (items within the chasing relation), and a dog-house, jeep, and plant (neutral items). The corresponding target scene on the right shows a lion chasing a horse (items within the chasing relation), a barn and soccer ball (neutral items), and a tiger (a distractor, featurally similar to the prompted tiger in the source scene). **Fig. 1B** shows an example of a reading *source* and *target* scene. The source scene on the left shows an elephant reading to a rabbit (items within the reading relation), and a tree, bench, and see-saw (neutral items). The corresponding target scene on the right shows a bird reading to a frog (items within the reading relation), a floatie and tent (neutral items), and a rabbit (a distractor, featurally similar to the prompted rabbit in the source scene). Stimuli were displayed on a 15-inch Dell laptop.

The directionality of relations within a pair of scenes was reversed to avoid children making choices based on spatial location alone. For example, in **Fig. 1A**, the direction of chasing is right to left in the source scene (the tiger on the right is chasing the boy on the left), whereas the direction of chasing is left to right in the target scene (the lion on the left is chasing the horse on the right).

2.2.2. Instruction stimuli

Four speech-alone and four speech + gesture instructional videos about how to solve scene analogies were created. All videos depicted chasing relations, but unlike the pre- and posttest problems, no item was circled. For both conditions, scenes were

¹ The sample size was limited due to the onset of the COVID-19 pandemic. While we are not able to collect more data for this study, we acknowledge that a larger sample would benefit the study. Working with this constraint, we have attempted to gain power in our analyses by conducting analyses at the trial level, while accounting for random effects of participants. However, we do acknowledge that replication of this study with a larger sample is needed.

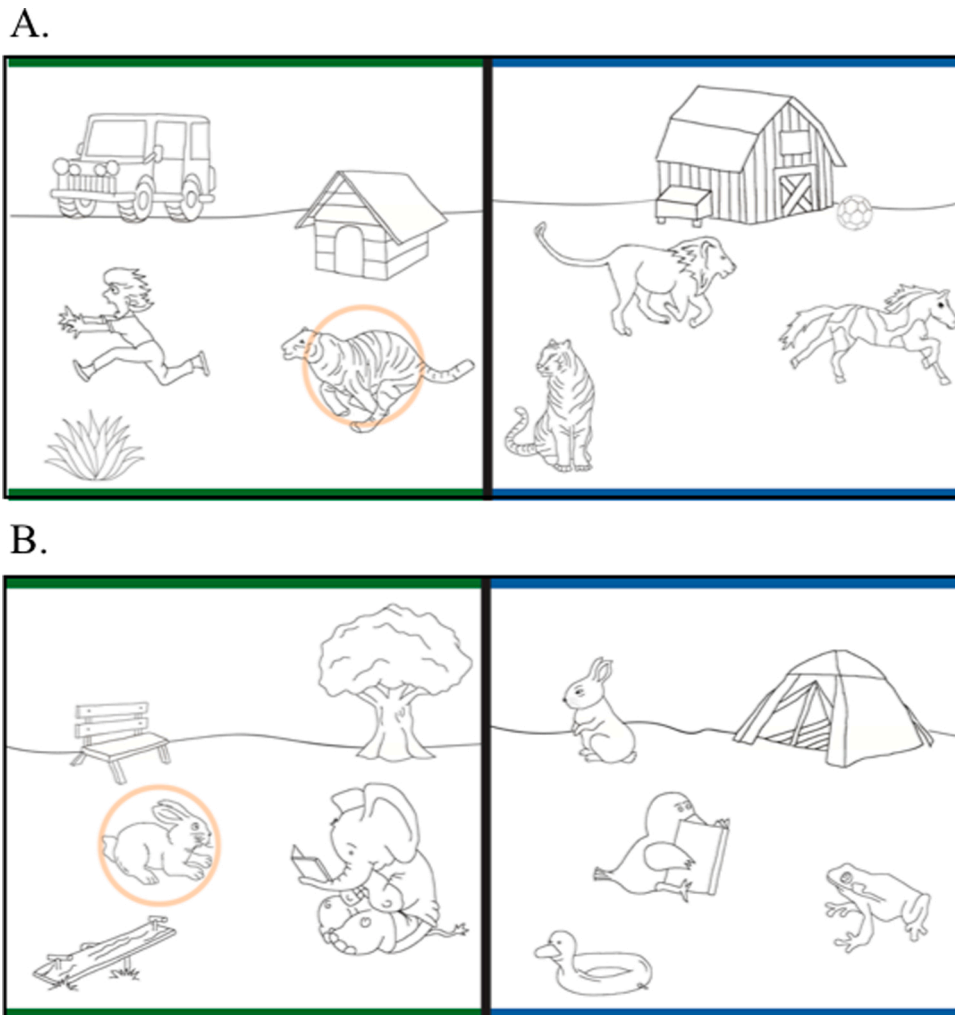


Fig. 1. A) Example trial of chasing relation category. B) Example trial of reading relation category.

accompanied by spoken instruction highlighting the relations between and within scenes. For example, “See, the boy is chasing the girl, and the dog is chasing the cat. This means the boy is in the same part of the pattern as the dog because they are both chasing, and the girl is in the same part of the pattern as the cat because they are both being chased”. In addition to spoken instruction, gestures were used to emphasize items and relations for children in the speech + gesture condition (see Fig. 2). In the example above, when the instructor said ‘The boy is chasing the girl’, a sweeping movement of the index finger traced a path from the boy to the girl, highlighting the chasing relation. The same sweeping gesture was used when the instructor said ‘... and the dog was chasing the cat’. Then, deictic gestures – pointing gestures used to indicate objects or locations – were used to simultaneously reference the items that were in the same parts of the relations. Items were indicated by a pointed index finger on each hand. When the instructor said, ‘This means the boy is in the same part of the pattern as the dog because they are both chasing’, simultaneous deictic gestures, made with the pointer fingers on each hand, pointed to the boy and the dog. Similarly, when the instructor said, ‘...and the girl is in the same part of the pattern as the cat because they are both being chased’, simultaneous deictic gestures pointed to the girl and the cat. Videos showed up to the entire hand depending on the item’s location.

To control for potential differences in inflection, the same audio track was used for the speech-alone and speech + gesture versions of each video. Videos were approximately 29 s.

2.2.3. Flanker inhibitory control and attention test

An NIH Toolbox version of the Flanker Inhibitory Control and Attention Test for ages 3–7 was administered to all participants. This task was a child-friendly version of a typical Erikson Flanker task, such that the children were asked to indicate which direction a cartoon fish was swimming. On some trials, a row of fish would be oriented in the same direction, including the target fish in the middle, signifying swimming in the same direction. On other trials, all fish except the target fish would be oriented in the same direction. The latter incongruent trials were the measure of IC. Inhibitory control score is calculated from the output provided by the NIH

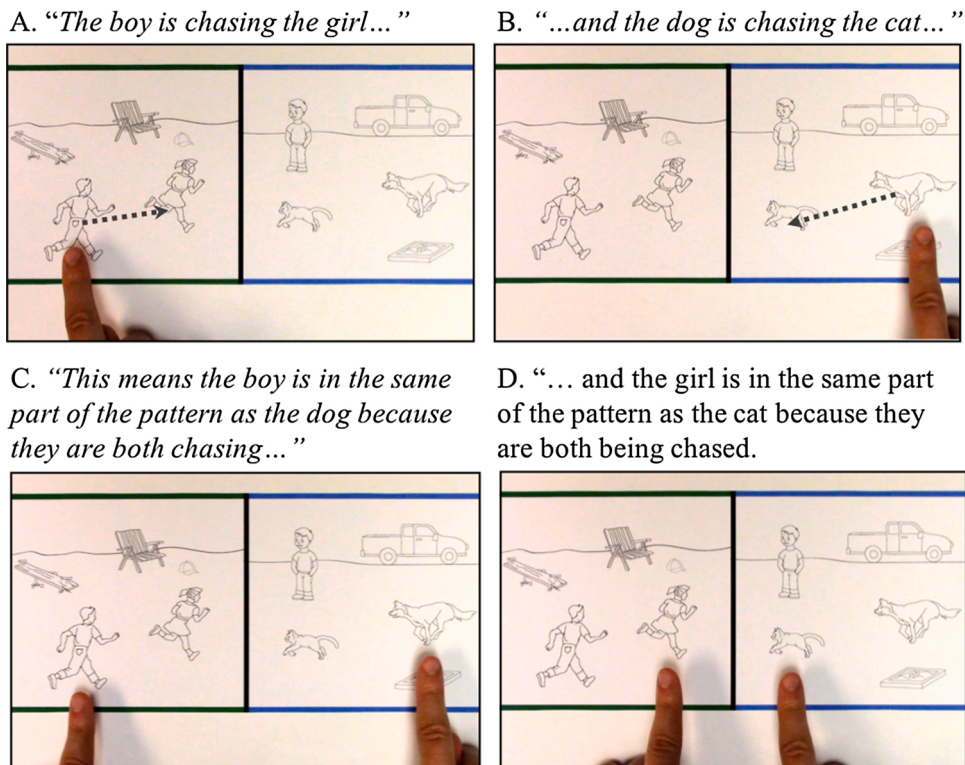


Fig. 2. Example instructional trial. Figure shows video stills from speech + gesture condition, and accompanying speech. Dotted arrows show the path the pointer finger took to show the chasing relation. Children heard identical speech, but did not see gesture, in the speech-alone condition.

toolbox software. Scores range from 0 to 10, where higher scores indicate mature IC.

2.2.4. Eye tracker

Eye tracking data were collected via corneal reflection using a Tobii 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 using standard animation blue dots. This step was followed by the collection and integration of gaze data with the presented instructional videos (described below) using Tobii Studio (Tobii Technology, Sweden). Data were extracted on the level of individual fixations as defined by Tobii Studio software—an algorithm determines if two points of gaze data are within a preset minimum distance from one another for a minimum of 100 msec, allowing for the exclusion of eye position information during saccades. After extraction, fixation location was queried at 8.33 msec intervals, to align with the native sampling frequency of the eye tracker (120 Hz).

2.3. Procedure

Children participated individually in a quiet setting at their school. Children were told they were going to play a picture game and completed a warm-up trial, which oriented them to the layout of test trials (i.e., two pictures with different colored borders; the terms left and right were not used because of the age of the children) and their task: For each set of scenes, their job was to find the pattern in the pictures. During the warm-up trial, the experimenter described the chasing relation, using language similar to the instructional videos, and asked the child to solve the relation. The explanation was repeated until the child chose the correct item. This introduction ensured that when children incorrectly answered a trial, it was not because they misunderstood the task. After completing the warm-up, children were seated approximately 40 cm in front of the laptop. Their position was calibrated and adjusted if necessary, and they were asked to remain as still as possible during the rest of the game, during which eye tracking data were collected.

Next, children completed a pretest. All children saw 12 pairs of scenes in a randomized order: six depicting the chasing relation, and six depicting the reading relation. An item in the source scene was circled and children were asked to identify the related item in the target scene: "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?". Responses were recorded for each trial and children were re-prompted if they did not respond.

After completing the pretest, children were asked to pay attention to instructional videos. Children were randomly assigned to watch four videos with speech-alone or speech + gesture instruction. In each video, children saw a source and target scene that each contained items in a chasing relation and neutral items. The target scene also contained a featural distractor. In both conditions, they heard spoken instruction explaining the relations within and between the scenes. The spoken instruction was accompanied by gesture

in the speech + gesture condition videos (see Materials for additional details). Following each video, children were presented with the scene-pair from the video with a source-scene item prompted, and asked to find the item in the picture with the blue edges that is in the same part of the pattern as the circled item. If an incorrect response was given, children were re-prompted until they selected the correct relational match to help emphasize the correct way to solve the scene analogies. Prompts to solve the instructional trials were only provided verbally for both conditions.

A posttest was administered after children watched the instructional videos. Children completed another set of 12 scene analogy problems (six chasing; six reading), identical in format to the pretest. Finally, children completed the NIH Toolbox Flanker Inhibitory Control and Attention Test for ages 3–7 to assess their IC.

Each part of the task was administered immediately following the previous part of the task. The entire procedure lasted 25–35 min (~10 min to familiarize the child with the eye tracker, for calibration, and to introduce the task, ~10 min for pretest, ~5 min for instruction, and ~10 min for posttest).

2.4. Data selection for eye tracking analyses

To address how visual attention related to children's IC and performance at pretest, and how visual attention was impacted by instruction at training and posttest, we considered the proportion of time spent looking at items within the scenes, and the degree to which children follow along with spoken instruction. Trials on which insufficient eye tracking data were collected (<65 % tracking during a trial) were excluded from analyses. If a child had insufficient tracking on greater than 75 % of the trials at a given timepoint (pretest, training, posttest), that child was excluded from analyses involving that timepoint of interest. This resulted in one child being excluded from analyses in which pretest was a timepoint of interest ($n_{\text{Speech-Alone}} = 1$) and three children being excluded from analyses in which posttest was a timepoint of interest ($n_{\text{Speech-Alone}} = 1$; $n_{\text{Speech+Gesture}} = 2$).

2.4.1. Fixation to the featural distractor

For pretest, training, and posttest trials, areas of interest (AOIs) were generated for each of the 10 items within the scene pairs using Tobii Studio, with a particular interest in how children visually attend to the featural distractor because previous research finds that children struggle to direct their attention away from featural matches within analogies (Guarino et al., 2019), and often choose the featural match rather than the correct, relational choice (e.g., Richland et al., 2006; see Fig. 1). The remaining spaces outside of these AOIs were collapsed into an "Other" AOI. To facilitate comparisons to previous work using other analogy formats (French & Thibaut, 2014; French, Gladly, & Thibaut, 2017; Gordon & Moser, 2007), proportion of time spent looking to each AOI was calculated by dividing the time looking to an AOI during the *latter third of solving time* by the total time looking during the latter third of solving time. We examined this portion of solving time, rather than initial solving time, because all children will likely attend to the source relation in the source scene to some degree initially given the saliency of the circled item, but the degree to which they continue to look to the featural match will impact their performance (Guarino et al., 2019). In contrast, during instruction, proportion of time spent looking to AOIs was calculated by dividing the time looking to an AOI throughout an *entire trial* by the total time looking during the entire trial. Proportion of time spent looking to AOIs were used as dependent measures in our analyses.

2.4.2. Following along during instruction

Because previous work suggests that gesture can help children follow along with spoken instruction and that this is predictive of learning (Wakefield, Novack et al., 2018), we also calculated a 'following score' for each training trial. Following scores were calculated by creating four time segments in which different relational comparisons were made by the instructor (see Fig. 2) and assessing whether children looked to AOIs highlighted in speech during each segment (i.e., during a given segment, children received a score of '1' if they looked to the relevant AOIs as they were labeled in speech and a '0' if they did not). Children could receive a score of 0–4 on each training trial. Scores calculated for each instructional trial were used in analyses.

3. Results

All analyses were conducted using R Studio (version 1.1.456), supported by R version 3.6.0. Analyses relied on the *lme4* package, which allows for mixed effects modeling (Bates, Maechler, Bolker, & Walker, 2015). When running mixed effects models through *lme4*, dummy coding was used, the default option for coding in this package. Appropriate reference levels for factors were assigned before each model was run: When condition was included in the models, the speech-alone condition was set as the baseline and compared against the speech + gesture condition. Because trials at pre- and posttest included two types of relations (reading and chasing), trial type was included as a predictor in all pre- and posttest models to account for any difference in performance between the two types of relations. Participant was included as a random factor in all models. An alpha level of .05 was used when evaluating statistical significance.

3.1. Do inhibitory control and visual attention predict behavioral performance?

Before considering our main questions of interest, we asked whether we could replicate previous findings showing that 4- and 5-year-old children perform poorly on scene analogy problems: They often choose the featural distractor rather than the correct relational choice (Richland et al., 2006), and this is associated with an inability to visually ignore the featural distractor (Guarino et al., 2019) and underdeveloped IC (2011, Dumas et al., 2018; Morrison et al., 2004; Viskontas et al., 2004). As expected, children

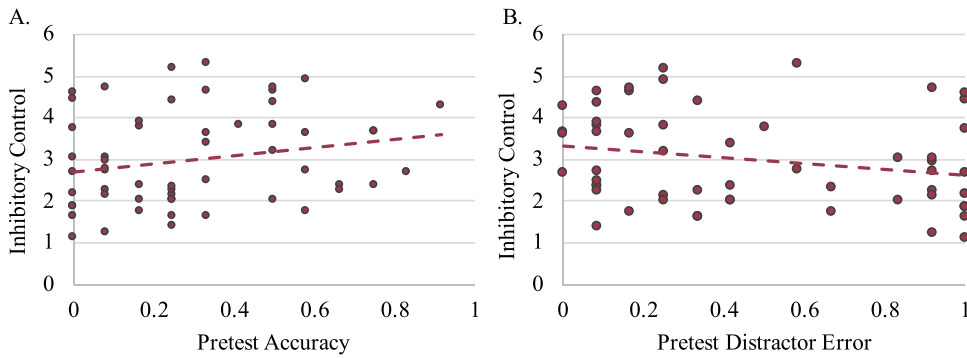


Fig. 3. A) Relation between inhibitory control and pretest accuracy. B) Relation between inhibitory control and pretest distractor error.

performed poorly at pretest, correctly answering 3.54 trials (out of 12; $SD = 3.04$) and choosing the featural distractor on nearly half of the trials ($M = 5.72$ out of 12 trials, $SD = 4.49$), allocated 26 % of their attention in the final third of the trial to the visual distractor, and showed variable but immature IC, scoring between 1.13 and 5.31 on the Eriksen Flanker task (scale range: 0–10).

To test whether there was a significant relation between children's likelihood to choose the correct choice or the incorrect featural distractor at pretest, and their IC and attention to the featural distractor, we constructed two binomial logistic regression models. In the first model, trial-level accuracy (0,1) served as the dependent measure with proportion looking to the distractor on a given trial and IC score as fixed factors. We found a main effect of proportion looking to the distractor, such that looking to the distractor negatively predicted accuracy ($\beta = -6.48$, $SE = 0.80$, $t = -8.06$, $p < .001$), but found no main effect of IC on accuracy ($\beta = 0.21$, $SE = 0.15$, $t = 1.37$, $p = .172$; Fig. 3).

In the second model, the same fixed factors were used but trial-level distractor error (0,1) served as the dependent measure. We found a significant main effect of proportion looking to the distractor, but this time in the opposite direction: Looking to the distractor was positively related to choosing the distractor ($\beta = 14.19$, $SE = 1.57$, $t = 9.05$, $p < .001$). In contrast to the first model, the main effect of IC was also significant, such that the lower a child's IC, the more likely they were to choose the distractor ($\beta = -0.76$, $SE = 0.38$, $t = -1.99$, $p = .047$). These findings suggest that looking to the distractor is associated with incorrectly choosing the distractor and failing to make the correct choice, but immature IC contributes more to children's bias toward featural matches than their accuracy.

3.2. Does children's inhibitory control relate to their visual attention patterns?

Given our conceptual replication of prior work demonstrating relations between IC and visual attention with children's analogical reasoning ability, our first main goal was to determine if one reason IC promotes success on analogical reasoning tasks is because it organizes visual attention away from featural distractors during problem solving. We constructed a generalized linear model with proportion of time spent looking to the distractor on each problem as the dependent measure, and children's Flanker IC score as the fixed factor. We found a marginal relation between IC and looking to the distractor, suggesting that the lower children's IC, the greater the proportion of time spent looking towards the distractor ($\beta = -0.04$, $SE = 0.02$, $t = -1.87$, $p = .068$; Fig. 4). Importantly, a similar model showed that IC did not predict whether children look at the distractor at least once during the entire trial ($\beta = -0.02$, $SE = 0.01$, $t = -1.42$, $p = .162$), suggesting that all children 'checked-in' with the distractor, but only those with lower IC continued to fixate on the distractor in the final third of solving time. Although the relation between IC and proportion looking to the featural distractor was trending and did not reach statistical significance, in combination with the finding that simply checking in with the featural distractor

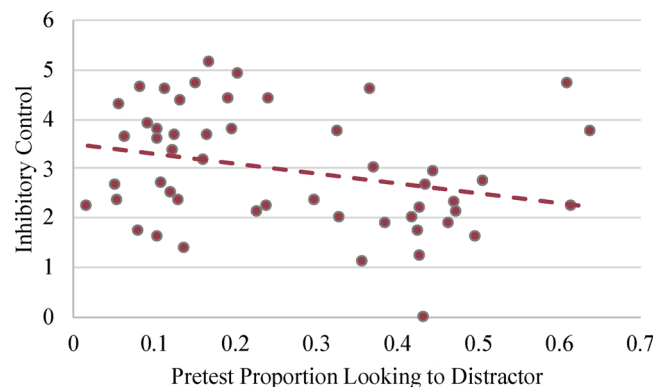


Fig. 4. Relation between inhibitory control and proportion of time spent looking to the distractor during the latter third of solving time.

was not at all related to IC, we suggest the issue of statistical significance may relate more to statistical power than a lack of relation. Still, we wish to emphasize that this result is *marginal*, and not robust evidence for a relation between IC and amount of looking to the distractor. While this is a relation that makes theoretical sense, it should be replicated in future studies with a larger number of participants.

3.3. Can including gesture in instruction facilitate children's analogical reasoning ability?

Our second goal was to ask whether gesture – a tool that has been shown to facilitate learning in many domains – can improve children's analogical reasoning ability, particularly by promoting beneficial looking patterns. To do this, we considered (1) whether behavioral performance at posttest differed between conditions, (2) whether visual attention patterns during instruction differed between conditions, (3) whether visual attention patterns during instruction predict posttest behavioral performance, and finally (4) whether the relation between visual attention at posttest changed due to instruction. This set of analyses addresses whether any behavioral changes at posttest, compared to pretest, are due to the effect of instruction on visual attention, either during instruction or following instruction. Before addressing these questions, we confirmed that there were no pretest differences between children who had been randomly assigned to each condition on behavioral performance measures (accuracy: $\beta = 0.22$, $SE = 0.42$, $t = 0.51$, $p = .607$; distractor error: $\beta = -0.38$, $SE = 0.70$, $t = -0.54$, $p = .590$) or proportion looking to the distractor ($\beta = -0.01$, $SE = 0.05$, $t = -0.16$, $p = .874$).

3.3.1. Does behavioral performance at posttest differ by condition?

First, we asked if accuracy was impacted by instruction. On average, children in the speech-alone condition correctly answered 5.07 problems (out of 12, $SD = 3.42$) and children in the speech + gesture condition correctly answered 4.48 problems ($SD = 3.16$). We constructed a binomial logistic regression model with accuracy (0,1) as the dependent measure, and condition (speech-alone, speech + gesture), timepoint (pretest, posttest), and two 2-way interactions (condition x timepoint; trial type x timepoint) as fixed effects. IC was not included in models predicting accuracy at posttest because there was not a pre-existing relation between IC and accuracy prior to instruction.

We found no evidence of an interaction between condition and timepoint predicting accuracy ($\beta = -0.42$, $SE = 0.27$, $t = -1.57$, $p = .116$), which would have suggested that one condition supported *greater* gains than the other condition, and no main effect of condition ($\beta = 0.22$, $SE = 0.42$, $t = 0.52$, $p = .604$). But, we did find evidence that children made performance gains from pretest to posttest: We found a significant interaction between trial type and timepoint ($\beta = 0.93$, $SE = 0.26$, $t = 3.53$, $p < .001$). Posthocs revealed that children significantly improved on both reading ($\beta = 0.58$, $SE = 0.27$, $t = 2.17$, $p = .029$) and chasing ($\beta = 1.03$, $SE = 0.27$, $t = 3.79$, $p < .001$) trials across timepoint, suggesting that improvement is seen on both trial types, although this effect was stronger for trials similar to those on which children received instruction. Together, these results suggest that all children learn from instruction, regardless of what type of instruction they received, and that children's performance improves both for trials similar to those used during training and those that require more generalization.

Next, we asked if choice of the featural distractor was impacted by instruction. On average, children in both conditions chose the distractor less at posttest than they did at pretest. Children in the speech-alone condition chose the distractor on 4.18 problems ($SD = 4.76$) compared to 6.07 problems ($SD = 4.99$) at pretest, and children in the speech + gesture condition chose the distractor on 4.59 problems ($SD = 3.74$), compared to 5.38 problems ($SD = 4.00$) at pretest. To test whether children showed significant decreases in distractor error from pretest to posttest, we constructed a similar binomial logistic regression model with trial level distractor error (0,1) as the dependent measure, and condition (speech-alone, speech + gesture), timepoint (pretest, instruction), IC score, and three 2-way interactions (condition x timepoint; IC x timepoint; trial type x timepoint) as fixed factors. Due to the significant relation between IC and choice of distractor prior to instruction, IC was included in models predicting distractor error.

We found a significant main effect of timepoint on distractor choice ($\beta = -2.55$, $SE = 0.54$, $t = -4.76$, $p < .001$), where children were less likely to select the featural distractor on posttest trials than on pretest trials. However, in contrast to the previous model, we did find a significant interaction between condition and timepoint ($\beta = 0.93$, $SE = 0.32$, $t = 2.91$, $p = .004$), with posthocs revealing a marginal effect of timepoint for children in the speech+gesture condition ($\beta = -0.98$, $SE = 0.57$, $t = -1.72$, $p = .085$) and a significant effect of timepoint for children in the speech-alone condition ($\beta = 0.00$, $SE = 0.00$, $t = -4.45$, $p < .001$). Additionally, in contrast to model predicting accuracy, there was no significant interaction between trial type and timepoint ($\beta = -0.36$, $SE = 0.30$, $t = -1.18$, $p = .236$). Within the context of the significant main effect of timepoint, these additional results suggest that children were less likely to choose the featural distractor on both reading and chasing trials at posttest as compared to pretest, but that overall this decrease was marginal for the speech+gesture condition.

When considering the impact of IC on distractor choice, we found a significant interaction between timepoint and IC ($\beta = 0.46$, $SE = 0.14$, $t = 3.27$, $p = .001$). This interaction suggests that the impact of IC changes when children solve problems before and after instruction. Posthocs revealed that the relation between IC and choice of distractor is only evident at pretest ($\beta = -0.66$, $SE = 0.32$, $t = -2.06$, $p = .040$) and not at posttest ($\beta = -0.14$, $SE = 0.37$, $t = -0.37$, $p = .711$), suggesting that instruction weakens the relation between IC and performance that was evident at pretest.

Together, these results suggest that whereas all children show improved accuracy after receiving instruction, the degree to which they improve does not differ based on receiving instruction with only speech or instruction with speech and gesture. However, while both conditions show a decrease in the likelihood of choosing the distractor after receiving instruction, the degree of change from pre- to posttest is different between conditions: The significant interaction between condition and timepoint predicting distractor choice, in concert with posthoc analyses showing a significant effect of timepoint for the speech-alone condition, but a marginal effect of

timepoint for speech + gesture condition, suggests that the decrease in distractor choice from pre- to posttest is *greater* for children in the speech-alone condition than children in the speech + gesture condition. Further, separate analyses show that the impact of IC is weakened by both forms of instruction, such that at posttest there is no longer a significant relation between children's IC and their analogical reasoning ability, suggesting that both types of instruction impact the role of children's existing IC in their ability to solve analogy problems.

3.3.2. Do visual attention patterns during instruction differ between conditions?

Although we found no significant differences in behavioral performance between conditions at posttest, we sought to understand whether gesture instruction might drive visual attention differently than speech-alone instruction during instruction, and whether such instruction could change how children subsequently solve and attend to problems at posttest. Specifically, we asked if instruction impacts two visual attention patterns: Proportion of time spent looking to the featural distractor (mirroring the visual attention patterns assessed at pre- and posttest) and ability to synchronize visual attention with the spoken instruction (following score).

To test whether children oriented their visual attention differently between conditions, we constructed two generalized linear models with visual attention measures on each problem (proportion of time spent looking to the distractor, following score) as the dependent measures, and children's IC score, condition, and average pretest accuracy as fixed factors. We found main effects of condition in both models: Children in the speech + gesture condition allocated significantly less of their attention to the distractor ($M = 0.06$, $SD = 0.01$) than children in the speech-alone condition ($M = 0.11$, $SD = 0.05$; $\beta = -0.05$, $SE = 0.01$, $t = -5.34$, $p < .001$). When considering children's ability to follow along with spoken instruction, children in the speech + gesture condition showed better alignment between their visual attention and spoken instruction than children in the speech-alone condition (speech + gesture: $M = 2.81$ following score, $SD = 0.90$); speech-alone: $M = 2.02$ following score, $SD = 0.96$; $\beta = 0.82$, $SE = 0.26$, $t = 3.15$, $p = .002$). These findings suggest that the inclusion of gesture helped orient children's attention to items as they were referenced in speech and drew attention away from the featural distractor, a visual attention pattern associated with higher IC. Interestingly, although our results at pretest suggest that IC marginally predicted looking to the distractor, there was no evidence of this relation during instruction ($\beta = -0.00$, $SE = 0.01$, $t = -0.62$, $p = .536$), suggesting that if this pattern did indeed exist during pretest, the presence of instruction was able to direct children's visual attention despite their IC. Finally, when considering relation between children's pretest performance and visual attention measures at instruction, we found that pretest performance predicted proportion of time looking to the distractor, such that children who scored higher on pretest looked less to the distractor ($\beta = -0.04$, $SE = 0.02$, $t = -2.75$, $p = .009$); however, pretest performance did not predict following instruction ($\beta = 0.50$, $SE = 0.45$, $t = 1.13$, $p = .266$) suggesting that instruction was enough to guide all children's attention to relationally important items in scene analogies.

3.3.3. Does visual attention during instruction predict subsequent performance?

Next, we asked if how children orient their visual attention during instruction impacts their posttest behavioral performance. Because we did not find differences between conditions in posttest performance, even though speech + gesture supports more effective visual attention during instruction, we asked if, across both conditions, visual attention during instruction impacts later analogical reasoning ability. We constructed a pair of binomial logistic regression models: one assessing the relation between looking to the distractor and choice of the distractor, and another assessing the relation between following score and accuracy. First, we constructed a model with trial-level distractor error at posttest (0,1) as the dependent measure and proportion looking to the distractor at instruction, IC score, and average pretest accuracy as fixed factors. We did not find a significant effect of IC ($\beta = 0.15$, $SE = 0.19$, $t = 0.76$, $p = .447$), supporting the previously discussed findings that IC no longer impacts posttest performance. And while we found that pretest accuracy related to posttest distractor error ($\beta = -8.62$, $SE = 1.05$, $t = -8.21$, $p < .001$), such that children who made fewer errors at pretest also made fewer errors at posttest, importantly, we did not find a relation between looking to the distractor during instruction and posttest distractor error ($\beta = 3.98$, $SE = 4.59$, $t = 0.87$, $p = .386$). Using a similar model, we asked whether children's ability to following along with spoken instruction related to their posttest accuracy. While we found a significant effect of posttest accuracy ($\beta = 5.86$, $SE = 0.48$, $t = 12.22$, $p < .001$), we did not find an effect of IC ($\beta = -0.01$, $SE = 0.09$, $t = -0.11$, $p = .915$) or following score ($\beta = 0.03$, $SE = 0.10$, $t = 0.33$, $p = .742$). Together, these findings suggest that while both types of instruction are able to weaken the relation between IC and performance, how children oriented their visual attention during instruction did not directly relate to their analogical reasoning ability during subsequent problem solving. In other words, while speech + gesture instruction can support effective visual attention patterns (i.e., less attention to the featural distractor and effective synchronization of visual attention with spoken instruction), this effect does not extend beyond the instructional session.

3.3.4. Does the relation between visual attention at posttest change due to instruction?

Finally, to further understand whether the relation between visual attention and performance changes due to instruction, we asked whether the relation between visual attention and behavioral performance evident at pretest changes at posttest. Recall that at pretest visual attention to the featural distractor negatively predicted accuracy and positively predicted choice of the distractor. These relations held at posttest: Using models similar to those used at pretest, we constructed two binomial logistic regression models. In the first model, trial-level accuracy (0,1) served as the dependent measure with proportion looking to the distractor, IC score, and average pretest accuracy were treated as fixed factors. In the second model, we asked if trial-level distractor error (0,1) was predicted by attention to the featural distractor. We found that the relation between attention to the distractor and accuracy seen at pretest remained at posttest ($\beta = -4.38$, $SE = 0.76$, $t = -5.74$, $p < .001$), such that children who attended more to the featural distractor were less likely to choose the correct, relational item. Similarly, the relation between attention to the distractor and distractor error remained at posttest ($\beta = 12.36$, $SE = 1.56$, $t = 7.94$, $p < .001$), suggesting that children who attended to the featural distractor more

made more errors at posttest. Together, these results suggest that while instruction supports effective visual attention in the moment, how children orient their visual attention still plays a large role in their analogical reasoning ability after instruction.

4. Discussion

The goals of the current study were to (1) examine how IC might support children's ability to solve scene analogies by directing visual attention away from featural distractors, and (2) determine whether providing instruction with gesture supports successful analogical reasoning, potentially by changing the way children allocate their attention during problem solving. Although marginally significant, our results are suggestive that IC promotes attention away from the featural distractor before instruction. At instruction, we see evidence that instruction with gesture can encourage similar patterns of visual attention: Children who received instruction through speech + gesture instruction were better able to follow-along with instruction that highlighted relational items, and allocated less attention to the featural distractor than children who received speech-alone instruction. However, these differences in visual attention patterns did not predict posttest performance: Children showed similar performance gains across the two conditions as well as gains in accuracy for both chasing and reading trials, suggesting that children were learning about relations in a general sense, not simply improving the representational strength of the specific chasing relation – suggesting the utility of this type of instruction for improving general, not problem-specific, understanding of analogy. And interestingly, children in the speech-alone condition made significantly fewer choices of the distractor, whereas this decrease was marginal for children in the speech + gesture condition. Finally, while the relation between children's visual attention and performance did not change due to instruction, the relation between children's IC and performance was weakened following both types of instruction, suggesting that instruction was able to overcome the pre-existing association between children's IC and their analogical reasoning ability. We situate these results within the analogical reasoning and gesture learning literature.

4.1. Understanding the relation between visual attention, IC, and behavioral performance

As in previous work, our results suggest that visual attention patterns relate to children's ability to solve analogies (e.g., Starr et al., 2018), and that children's low IC hinders their performance (e.g., Richland et al., 2006). We also find support for a prediction made by Simms and colleagues (2018), that the Eriksen Flanker task is related to analogical reasoning ability in young children.

But more importantly, we move beyond considering the relation of behavioral performance with visual attention or IC, and consider how all three of these measures relate to one another. Whereas researchers have inferred that improved reasoning skills *should* be associated with ability to direct attention away from superficial matches (e.g., Richland et al., 2006); here, we provide direct evidence for this claim. We show that when children are solving analogical reasoning problems before having any instructional support, allocation of attention to the featural distractor is systematically related to analogical reasoning performance and marginally related to our measure of IC. We also find that children's IC systematically related to their choice of the featural match prior to instruction, which provides direct evidence of a relation between children's IC and analogical reasoning ability.

These results have important implications for future research on the development of analogical reasoning ability. Researchers studying how individual differences impact the development of analogical reasoning have traditionally employed separate behavioral measures of IC like the flanker task, which are collected before or after children complete an analogical reasoning task. From a practical perspective, our work shows that measuring visual attention during problem solving could be used as a more direct proxy for assessing a child's IC than other behavioral measures. While replication of the current findings are warranted, future research could even use visual attention as a behavioral measure of children's IC *during* tasks of analogical reasoning. This would be particularly informative prior to any kind of instructional intervention, where IC could serve as an index of children's solving strategy and reasoning ability. Our results are also interesting from a theoretical perspective. Whereas prior work suggests that IC is one contributing component to the development of children's analogical reasoning ability, *how* exactly IC functions during problem solving is unclear. Here, our results provide further support for the argument that visual attention strategies are related to children's IC.

However, when taken as a whole, our results suggest a more complicated story about the relation between IC (as measured by the Eriksen flanker task) and behavioral performance. Our measure of IC correlated with children's choice of the featural distractor at pretest, but this relation was *not* found at posttest. This suggests that either type of instruction, much like relational learning (Richland et al., 2010), can help to compensate for children's lack of IC at critical moments of development when IC is particularly important for determining analogical reasoning ability (e.g., Doumas et al., 2018). The fact that a short instructional period can disrupt the relation between the flanker task and behavioral performance following instruction may indicate that the flanker task is capturing an aspect of IC that has more to do with cognitive control, rather than avoiding featural lures. The instruction may have acted as a guide to the goal of the task and provided a tool to structure children's thought process at posttest that was not available during pretest.

In contrast, whereas IC was systematically related to choice of the featural distractor at pretest but not at posttest, visual attention *did* remain systematically related to behavioral performance across time points. When children allocated more attention to distractor, they were more likely to choose the distractor. We suggest that these visual attention patterns may reflect the aspect of IC that acts as a gating mechanism and that either a longer intervention or maturation may be required to change these patterns. Additional work must be conducted to further tease apart these different aspects of IC, and how they relate to visual attention and behavioral performance.

4.2. Understanding the impact of gesture during problem solving

In addressing our second question, we found that when viewing instruction that incorporated gesture, children allocated their

visual attention differently than when viewing instruction without gesture. Gesture promoted the mature looking pattern that had been associated with higher IC at pretest – directing attention away from the featural distractor. We also found, in line with previous research on mathematical equivalence instruction (Wakefield, Novack et al., 2018), that gesture instruction supported children's ability to follow along with spoken instruction. That is, children were more likely to attend to the relational items in the scenes as they were referenced in speech if they received speech + gesture instruction, compared to speech-alone instruction. This impact of gesture is likely due to its ability to direct visual attention, but also its ability to facilitate comparison processes. Previous work has demonstrated that gestures can physically embody links between representations, which can serve to disambiguate spoken instruction. For example, Wakefield, Novack et al. (2018) found that when the connection between an instructor's words and the physical environment is unclear during mathematical instruction, gesture is able to facilitate the link between words and what they map on to, and thereby, disambiguate spoken referents. Grassmann and Tomasello (2010) made similar conclusions: Pointing gestures help young children disambiguate complex or contradictory verbal instructions in order to retrieve an object. Additionally, previous research suggests that seeing gesture should help to alleviate limitations of processing instruction through an auditory channel alone: Learning is facilitated when information is presented through speech and gesture simultaneously, rather than sequentially (Congdon et al., 2017).

While we find that gesture supports effective visual attention patterns during the instructional session, we did not find that those visual attention patterns predict posttest performance or visual attention patterns, indicating that gesture can encourage looking patterns associated with high IC when it is being used, but that these effects are transient and do not create a lasting change in solving strategies. The finding that speech + gesture instruction facilitates learning to the same degree as speech-alone instruction is not novel: Previous work within the gesture-for-learning literature has also found that there are contexts in which gesture does not always support learning (e.g., Congdon et al., 2018; Guarino & Wakefield, 2020; Post, Van Gog, Paas, & Zwaan, 2013; Wakefield & James, 2015). Both the form of the gesture itself (Dargue & Sweller, 2018a, 2018b) and the characteristics of the learner, such as their degree of prior knowledge within a given domain (mathematics: Congdon et al., 2018; grammar: Post et al., 2013; word learning: Wakefield & James, 2015), can impact whether or not gesture supports learning. For example, Post et al. (2013) found that when asking children to learn grammatical rules, only children with higher levels of general language skills benefited from the inclusion of gesture during training. They suggested that these findings were due to an expertise reversal effect, such that for children with lower levels of language skills the added cognitive effort needed to process gesture along with verbal instruction actually hurt their ability to learn. This is just one example in a line of work that has demonstrated that gesture's effects can depend on the context in which it is learned, including the characteristics of the learner.

However, these findings do diverge from much of the previous literature which finds that children benefit more from instruction with gesture than instruction with speech alone in a number of domains (e.g., mathematical equivalence: Congdon et al., 2017; Wakefield, Novack et al., 2018, measurement: Congdon et al., 2018, word learning: Wakefield & James, 2015; Wakefield, Hall et al., 2018, spatial-task domains: Chu & Kita, 2008; Valenzano et al., 2003). Why did children show no advantage when instructed through speech + gesture versus speech-alone, especially given that gesture *did* encourage visual attention patterns that should be helpful for problem solving? The discrepancy between our current work and previous literature may be a consequence of 1) the domain being taught and 2) how gesture both directs and constrains attention to the problem space and how this relates to what a child must learn.

When children are instructed with speech + gesture versus speech-alone in mathematical equivalence tasks, children already have some degree of prior experience with arithmetic concepts that sets the stage for mastering math equivalence, such as the ability to add single-digit numbers together. What needs to be learned is the meaning of the equal sign. In other words, children have the basic tools that are necessary to support correct problem solving, but must overcome a fundamental misunderstanding of *how* to solve the problem. Similarly, in word learning studies, young children have already mastered the ability to map new words to referents and are proficient language learners; gesture is there to help support the new connections between novel words and referents that they are asked to make during the study session. In contrast, when considering analogical reasoning, children have *not* fully acquired the cognitive abilities that are fundamental to solving analogies, struggling with immature inhibitory control and working memory. They must also work against biases to focus on featural properties when learning, which were crucial in early childhood for noun learning (e.g., Murphy et al., 2021). It may be that in this case, gesture is *not* able to give children the extra boost that is often evident in the gesture-for-learning literature. The present study therefore extends the line of research addressing gesture's impact on learners to demonstrate that there are situational factors, such as the learning domain, that can impact whether or not using gesture along with spoken instruction benefits the learner beyond speech alone instruction.

Beyond the domain, we can consider a more nuanced difference in the current learning paradigm and those used in previous studies showing greater gains from gesture: How gesture interacts with what a child must learn is important – or unimportant – when solving a problem. One of the puzzling findings in the current study was that children in both conditions made similar, significant gains in accuracy from pre- to posttest, but children in the speech-alone condition also made significantly fewer choices of the distractor, whereas this result was in the same direction, but marginal, for children in the speech + gesture condition. Why was gesture not as helpful as speech alone in decreasing the choice of the distractor, when it actually *directed children away from* the featural distractor during instruction? It may be that this feature of gesture – its ability to direct visual attention – is actually a double-edged sword. In directing a child's focus, gesture also *constrains* this focus, which may not allow for necessary exploration of the problem space. In other words, in the context of a scene analogy problem, gesture may interfere with the likelihood a child will become familiar with the featural distractor and its irrelevance for solving a scene analogy problem. This may also explain why following ability did not predict posttest performance. In contrast, if we consider how gesture is used when teaching the concept of mathematical equivalence, what a child has to learn is inherent in the structure of the equation – there is no additional information present that may help a child make sense of equivalence, beyond the components of the equation. Gesture is great at highlighting these problem components, and so the ability to direct and constrain attention has no drawbacks in this learning context. Taken together, we suggest that gesture is great at

orienting visual attention, but that that may not always be a positive: For some learning contexts, it may be necessary to explore the problem space in order to fully understand what is and is not relevant for successful problem solving. This possibility could be tested in future studies, asking whether highlighting the distractor with gesture, and incorporating *why* it is not important when solving analogy problems into both spoken and gestured instruction.

Returning to our previous discussion of IC, it may be that either form of instruction can promote better cognitive control for children who are trying to solve analogical reasoning problems – which would lead to the decoupling of the relation between the Eriksen Flanker measure of IC and behavioral performance at posttest. Beyond the potential impact of instruction on children's cognitive control, it may be that both forms of instruction were able to strengthen children's relational knowledge representations. Previous computational simulations of analogical reasoning development have found that IC is not only related to one's ability to manage distraction, but also their ability to form new relational representations (Doumas et al., 2018; Morrison et al., 2011). Doumas et al. (2018) tested a model which finds through simulations that higher IC may allow for increased growth in relational knowledge representations, which supports more effective relational reasoning. Therefore, it may be that both forms of instruction in the present study were able to facilitate building relational representations of the task strategy by reducing IC demands to some extent. While the association between inhibitory control and building representations, or rule-learning, is well-established (e.g., Blackwell, Chatham, Wiseheart, & Munakata, 2014; Chevalier & Blaye, 2008; Diamond et al., 2002, Diamond, Kirkham, & Amsos, 2002; Egner & Hirsh, 2005), future work with this paradigm should further investigate whether training knowledge representations reduces children and adults' IC demands during the reasoning process.

However, despite the successful decoupling of IC and analogical reasoning performance, the gating mechanism component of IC is likely more maturational and thereby more difficult to affect. While gesture instruction may be able to direct attention effectively during the training session, perhaps children cannot overcome the maturational limitations of the gating mechanism of IC to sustain these advantageous patterns of visual attention when instructional support is no longer provided at posttest. It may be that children who are older than those in the present study and have that gating mechanism in place, but do not understand what rule should be applied when gating information into working memory, may benefit *more* from instruction than those in the current study. And, it may be that older children *would* show more benefit from gesture instruction than speech alone instruction. This would align with previous work that suggests gesture is a more advantageous tool for children with a certain level of prior knowledge or ability related to the to-be-learned concept (Congdon et al., 2018; Wakefield & James, 2015). Looking across a wider age range, future work could consider this possibility.

5. Conclusion

Because analogical reasoning is important for a host of skills necessary for future academic achievement and career success (e.g., innovation, creativity, inductive reasoning; Gentner, 2010), we aimed to better understand *why* children struggle with this ability and *how* we can support early development of these skills. Specifically, we aimed to better understand how visual attention might relate to IC, a factor that has been posited to contribute to the protracted development of analogical reasoning, and to ask whether gesture could support analogical reasoning ability. Our results extend previously established associations between IC and children's analogical reasoning by identifying a relation between IC and visual attention measures. Surprisingly, we found that whereas gesture is effective for directing visual attention in a beneficial way, children show equal gains in accuracy after speech-alone and speech + gesture instruction, and surprisingly, they show a significant decrease in choice of distractor after speech-alone instruction, whereas this change is marginal following speech + gesture instruction. These findings suggest that our instruction did have an impact on learning, but also raise important questions about the general claim that gesture will boost learning above-and-beyond speech-alone instruction, and highlight a need to further consider how aspects like domain and the way gesture interacts with what must be learned affect the utility of gesture as a teaching tool. Further, the finding that both analogical reasoning *and* the relation between IC and analogical reasoning ability is malleable to instruction provides insight into the developmental process of relational reasoning, revealing that it is not only driven by either maturation or knowledge acquisition, but also may be impacted by directed socialization.

CRedit authorship contribution statement

Katharine F. Guarino: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Elizabeth M. Wakefield:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - original draft, Writing - review & editing. **Robert G. Morrison:** Conceptualization, Methodology, Resources, Writing - review & editing. **Lindsey E. Richland:** Conceptualization, Methodology, Resources, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgements

Funding for this study was provided by Loyola University Chicago to E. Wakefield. The authors would like to thank Erick Bautista, Natalie Schomas, Sarah Heimberg, Madeline Jurcev, and Sasha Stojanovich for their help with data collection and coding, and

children, families and schools for their participation.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.cogdev.2021.101040>.

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67. <https://doi.org/10.18637/jss.v067.i01>
- Begolli, K. N., Richland, L. E., Jaeggi, S. M., Lyons, M., Klostermann, E. C., & Matlen, B. J. (2018). Executive function in learning mathematics by comparison: Incorporating everyday classrooms into the science of learning. *Thinking & Reasoning*, 24, 1–34. <https://doi.org/10.1080/13546783.2018.1429306>
- Blackwell, K. A., Chatham, C. H., Wiseheart, M., & Munakata, Y. (2014). A developmental window into trade-offs in executive function: The case of task switching versus response inhibition in 6-year-olds. *Neuropsychologia*, 62, 356–364. <https://doi.org/10.1016/j.neuropsychologia.2014.04.016>
- Chevalier, N., & Blaye, A. (2008). Cognitive flexibility in preschoolers: The role of representation activation and maintenance. *Developmental Science*, 11, 339–353. <https://doi.org/10.1111/j.1467-7687.2008.00679.x>
- Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology General*, 137, 706–723. <https://doi.org/10.1037/a0013157>
- Congdon, E. L., Kwon, M. K., & Levine, S. C. (2018). Learning to measure through action and gesture: Children's prior knowledge matters. *Cognition*, 180, 182–190. <https://doi.org/10.1016/j.cognition.2018.07.002>
- Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction*, 50, 65–74. <https://doi.org/10.1016/j.learninstruc.2017.03.005>
- Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, 84, 1863–1871. <https://doi.org/10.1111/cdev.12097>
- Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, 106, 1047–1058. <https://doi.org/10.1016/j.cognition.2007.04.010>
- Dargue, N., & Sweller, N. (2018a). Donald Duck's garden: The effects of observing iconic reinforcing and contradictory gestures on narrative comprehension. *Journal of Experimental Child Psychology*, 175, 96–107. <https://doi.org/10.1016/j.jecp.2018.06.004>
- Dargue, N., & Sweller, N. (2018b). Not all gestures are created equal: The effects of typical and atypical iconic gestures on narrative comprehension. *Journal of Nonverbal Behavior*, 42, 327–345. <https://doi.org/10.1007/s10919-018-0278-3>
- Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38, 352–362. <https://doi.org/10.1037/0012-1649.38.3.352>
- 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>
- Dunbar, K. N., & Klahr, D. (2012). Scientific thinking and reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *Oxford handbook of thinking and reasoning*. New York: Oxford University Press.
- Egner, T., & Hirsh, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8, 1784–1790. <https://doi.org/10.1038/nn1594>
- Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development*, 86, 1386–1405. <https://doi.org/10.1111/cdev.12381>
- French, R. M., & Thibaut, J. P. (2014). Using eye-tracking to predict children's success or failure on analogy tasks. *Livre/Conférence Proceedings of the Thirty-Sixth Annual Meeting of the Cognitive Science Society*, 2222–2227.
- French, R. M., Glady, Y., & Thibaut, J. P. (2017). An evaluation of scanpath-comparison and machine-learning classification algorithms used to study the dynamics of analogy making. *Behavior Research Methods*, 49, 1291–1302. <https://doi.org/10.3758/s13428-016-0788-z>
- 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., & Smith, L. A. (2013). Analogical learning and reasoning. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 668–681). New York, NY: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195376746.013.0042>
- 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>
- Goswami, U. (2001). Analogical reasoning in children. *Proceedings of the International Conference on Conceptual Structures*, 437–470. <https://doi.org/10.4324/9781315804729>
- Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, 13, 252–263. <https://doi.org/10.1111/j.1467-7687.2009.00871.x>
- Guarino, K. F., & Wakefield, E. M. (2020). Teaching analogical reasoning with co-speech gesture shows children where to look, but only boosts learning for some. *Frontiers in Psychology*, 11, 1–14. <https://doi.org/10.3389/fpsyg.2020.575628>
- Guarino, K. F., Wakefield, E. M., Morrison, R. G., & Richland, L. E. (2019). Looking patterns during analogical reasoning: Generalizable or task-specific? *Forty-First Annual Meeting of the Cognitive Science Society*, 387–392.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- 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.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: The University of Chicago.
- 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>
- Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., et al. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience*, 12, 260–271. <https://doi.org/10.1162/089892904322984553>
- Murphy, A., Zheng, Y., Shivaram, A., Vollman, E., & Richland, L. E. (2021). Bias and sensitivity to task constraints in spontaneous relational attention. *Journal of Experimental Child Psychology*, 202, 1–23. <https://doi.org/10.1016/j.jecp.2020.104981>
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Child Psychology*, 131, 5–15. <https://doi.org/10.1037/0096-3445.131.1.5>
- Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science*, 25, 902–910. <https://doi.org/10.1177/0956797613518351>
- Ping, R., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, 44, 1277–1287. <https://doi.org/10.1037/0012-1649.44.5.1277>

- Post, L. S., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). Effects of simultaneously observing and making gestures while studying grammar animations on cognitive load and learning. *Computers in Human Behavior*, *29*, 1450–1455. <https://doi.org/10.1016/j.chb.2013.01.005>
- 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](https://doi.org/10.1016/S0885-2014(98)90003-X)
- Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science*, *24*, 87–92. <https://doi.org/10.1177/0956797612450883>
- Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews Cognitive Science*, *6*, 177–192. <https://doi.org/10.1002/wcs.1336>
- Richland, L. E., Chan, T.-K., Morrison, R. G., & Au, T. K.-F. (2010). Young children's analogical reasoning across cultures: Similarities and differences. *Journal of Experimental Child Psychology*, *105*, 146–153. <https://doi.org/10.1016/j.jecp.2009.08.003>
- 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>
- Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*, 426–435. <https://doi.org/10.1111/j.1467-7687.2012.01139.x>
- Rueda, M. R., Fan, J., Halparin, J., Gruber, D., Lercari, L. P., Mc-Candliss, B., et al. (2004). Development of attention during childhood. *Neuropsychologia*, *42*, 1029–1040. <https://doi.org/10.1016/j.neuropsychologia.2003.12.012>
- Simms, N., Frausel, R., & Richland, L. R. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology*, *166*, 160–177. <https://doi.org/10.1016/j.jecp.2017.08.005>
- 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. (Ed.). (1988). *The nature of creativity: Contemporary psychological perspectives*. New York, NY: Cambridge University Press.
- 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>
- 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>
- Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187–204. [https://doi.org/10.1016/S0361-476X\(02\)00007-3](https://doi.org/10.1016/S0361-476X(02)00007-3)
- 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. <https://doi.org/10.1037/0882-7974.19.4.581>
- Wakefield, E. M., & James, K. H. (2015). Effects of learning with gesture on children's understanding of a new language concept. *Developmental Psychology*, *51*, 1105–1114. <https://doi.org/10.1037/a0039471>
- Wakefield, E. M., Hall, C., James, K. H., & Goldin-Meadow, S. (2018). Gesture for generalization: Gesture facilitates flexible learning of words for actions on objects. *Developmental Science*, *21*. <https://doi.org/10.1111/desc.12656>
- Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*, *21*. <https://doi.org/10.1111/desc.12664>