

# Visual routines for extracting magnitude relations

**Audrey L. Michal, David Uttal, Priti Shah & Steven L. Franconeri**

**Psychonomic Bulletin & Review**

ISSN 1069-9384

Psychon Bull Rev

DOI 10.3758/s13423-016-1047-0



**Your article is protected by copyright and all rights are held exclusively by Psychonomic Society, Inc.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Visual routines for extracting magnitude relations

Audrey L. Michal<sup>1</sup> · David Uttal<sup>1</sup> · Priti Shah<sup>2</sup> · Steven L. Franconeri<sup>1</sup>

© Psychonomic Society, Inc. 2016

**Abstract** Linking relations described in text with relations in visualizations is often difficult. We used eye tracking to measure the optimal way to extract such relations in graphs, college students, and young children (6- and 8-year-olds). Participants compared relational statements (“Are there more blueberries than oranges?”) with simple graphs, and two systematic patterns emerged: eye movements that followed the verbal order of the question (inspecting the “blueberry” value first) versus those that followed a left-first bias (regardless of the left value’s identity). Question-order patterns led substantially to faster responses and increased in prevalence with age, whereas the left-first pattern led to far slower responses and was the dominant strategy for younger children. We argue that the optimal way to verify a verbally expressed relation’s consistency with visualization is for the eyes to mimic the verbal ordering but that this strategy requires executive control and coordination with language.

**Keywords** Eye movements and visual attention · Relations · Visualizations

**Electronic supplementary material** The online version of this article (doi:10.3758/s13423-016-1047-0) contains supplementary material, which is available to authorized users.

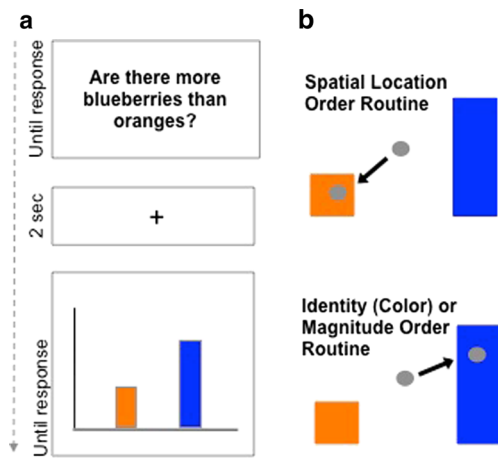
✉ Audrey L. Michal  
audrey.lustig.michal@northwestern.edu

<sup>1</sup> Department of Psychology, Northwestern University, 2029 Sheridan Rd, Evanston, IL 60208, USA

<sup>2</sup> Department of Psychology, University of Michigan, Evanston, IL, USA

Relational comparison is fundamental to higher-level cognition and consequently has received substantial attention in research (Halford, Wilson, & Philips, 2010; Hummel & Holyoak, 2005; Gentner, 1983). Many relational judgments require comparison of multiple types of relational representations, including visuals and text (Ainsworth, 2006; Bergey, Cromley & Newcomb, 2015). Assembling furniture requires integrating a diagram with verbal instructions, and solving a mathematical word problem can require matching verbal relations to a drawing. More generally, a fundamental part of scientific reasoning is the ability to compare evidence with theories (Kuhn & Pearsall, 2000), and the former is typically expressed through visualizations and the latter through verbal descriptions. Although all of these comparisons require people to compare particular visual relations to particular verbalizations, there is strikingly little work on how our visual system accomplishes this task. We explored the case study comparing relations in visually presented data in *graphs* to verbally described relations.

Extracting visual relations is surprisingly complex. Take the minimally simple bar graph in Fig. 1. At first glance, it contains a single relation between two objects. Closer inspection reveals far more. The two bars contain three dimensions that could be important: magnitudes (small and large), colors (blue and orange), and spatial locations (left and right). Among these six values, there are a surprising number of potential relations that a graph reader might extract (Table 1). A viewer could focus on magnitude relations based on color (e.g., the blue bar is larger), color relations based on magnitudes (e.g., the larger bar is blue), or magnitude relations based on spatial locations (e.g., the left bar is smaller). Table 1 shows the 6 total variations in these descriptions, which doubles to 12 when each of these descriptions is inverted (e.g., the last example could instead state that the right bar is larger). The number doubles again to 24 when each is rephrased as a



**Fig. 1** (a) Example trial sequence. Participants first viewed a question about a magnitude relation in an upcoming graph, followed by a brief delay and a two-bar graph. The colors of the bars in the graph represented different types of fruits (e.g., orange = oranges, blue = blueberries). (b) Schematics showing three possible visual routines during graph viewing: *spatial location* (i.e., first saccade towards left bar; top), *identity or magnitude* (i.e., first saccade toward bar associated with the relational target of the question or toward the taller bar; bottom). Gray dots represent eye fixations

question, such as “which color is farther to the right?” Our graph only has *two bars*. It is no wonder that younger students have such difficulty dealing with graphs (Friel et al., 2001) and why advanced students still struggle with more complex graphs. Even experienced graph readers find graph comprehension difficult, taking time and effort equivalent to reading a paragraph (Carpenter & Shah, 1998).

Addressing the challenges of comparing visual and verbal relations requires that we understand how the expert viewer is able to choose between alternative interpretations of a visual relation. We argue that the many different ways of interpreting relations in a graph are not simply different ways of *thinking* about patterns in the graph but different ways of *seeing* those patterns. Specifically, we argue that the *visual routine* used by a viewer (Ullman, 1984)—the sequence of a graph’s values

that they attend to over time—constructs the relational expressions that they see. We use eye tracking to determine the routines that most efficiently extract a given relational description. We report the results of a set of experiments that reveal the presence and importance of these routines and their development across three age groups: 6-year-olds, 8-year-olds, and college students.

Seeing different descriptions of relations within a graph is likely a serial visual routine, because its more basic subprocess—perceiving categorical visual relations—seems to be a serial process. Finding a pair of objects in a particular relation (e.g., a plus above a minus sign) within a display of oppositely arranged pairs is tediously slow, perhaps requiring inspection of single pairs at a time (Logan 1994, 1995; Wolfe, 2001). Recent work suggests an even tighter capacity limit for relational processing, suggesting that we must inspect each individual object within a pair to overcome the visual system’s severe limitations in matching objects with their properties (Franconeri et al., 2012; Luck & Ford, 1998; see Hummel & Biederman, 1992, for a more optimistic proposal). Electrophysiological data confirms this prediction: when people judge relations (do you see a green object to the left of a red object, or vice-versa?), they systematically shift their attention to isolate one of the objects, suggesting that moving the eyes left may produce the relational description “green on the left,” whereas moving to the right may produce “red on the right” (Franconeri et al., 2012).

Past work on scene perception is consistent with a close link between the focus of attention and relational descriptions (Griffin & Bock, 2000; Papafragou et al., 2008). Cueing attention to a given object can change the way that people describe its relation within a scene. In a scene depicting a dog and a man running, people were more likely to describe it as “The dog chases the man” when the dog was subliminally cued and as “The man flees the dog” when the man was cued (Gleitman et al., 2007; Roth & Franconeri, 2012). In another study, people were faster to verify that an image contained “a star above a plus” compared with “a plus below a star” despite

**Table 1** Twelve descriptions of possible relations present in the graph in Fig. 1. The descriptions are based on all possible combinations of the three dimensions (color, magnitude, location) of the two bars. Framing each as a question doubles the number of potential candidates

Descriptor	Dimension		
	Color	Magnitude	Location
Orange	–	The orange bar is smaller	The orange bar is on the left
Blue	–	The blue bar is larger	The blue bar is on the right
Smaller	The smaller bar is orange	–	The smaller bar is on the left
Larger	The larger bar is blue	–	The larger bar is on the left
Left	The left bar is orange	The left bar is smaller	–
Right	The right bar is blue	The right bar is larger	–

the logical equivalence of these two statements (Clark & Chase, 1972). This preference for framing a relation as “above” seemed to be due to an attentional bias for the top object (which is common in other tasks, see Xu & Franconeri, 2015); the effect weakened under instructions to attend to the bottom.

These results are all consistent with the idea that the way that a viewer isolates objects with attention over time can control the relational description that they extract from a scene. For the graph in Fig. 1, when asked, “Are there more blueberries than oranges?”—what’s the optimal way to move attention to answer the question? There are three salient types of routines:

- *Magnitude*: Inspect the larger bar to verify whether it is blue. If this is the routine used, we would predict that the viewer should seek the larger bar, because seeking affirmations of a sentence is more efficient than seeking negations (Clark & Chase, 1972), and the sentence relies on the word “more” as opposed to “fewer.”
- *Sentence ordering of object identity (color)*: Inspect the blue bar and verify that it is larger. We would predict that in this case the viewer should seek the blue bar, because comparing the first-mentioned target (“blueberries”) of the visually depicted relation is more efficient than comparing the reference (“oranges”) object (Roth & Franconeri, 2012, for review).
- *Spatial Location*: Inspect the left bar and verify that it is smaller. We would predict that expert graph viewers should not use this strategy, because the x-axis is categorical, not ordinal, for graphs in the current study; thus, the spatial locations of the bars should be irrelevant (although they might be relevant for graphs where the x-axis indicated time or another ordered dimension). Despite its inefficiency, some viewers might use this strategy with a leftward bias, based on the pervasiveness of the left bias in scene-viewing and English reading order (Dickinson & Intraub, 2009; Tversky et al., 1991). The left bias even appears to extend to graph reading. In past work using a task similar to ours, participants were fastest to verify that a graph contained a certain trend, e.g., “A is greater than B,” when the left-to-right order of the bars matched the order of the sentence (Feeney & Webber, 2003).

We tested for the presence and effectiveness of these three routines using displays similar to those in Fig. 1. We used eye tracking as a proxy for the locus of visual attention, measuring the direction of the first saccade to either bar, and we used response time to simple questions about magnitude relations, e.g., “Are there more blueberries than oranges?” as a measure of the effectiveness of the routine. We predicted that efficient graph reading would reflect the *magnitude* or *identity* routines and that the *spatial location* routine should lead to poor performance.

## Experiment 1

### Method

**Participants** We collected data from 17 adults (14 females, mean age = 23 years) with normal or corrected-to-normal vision. Participants were compensated \$10. Data from one participant was discarded, because eye movements occurred on only 25 % of trials. Sample sizes were determined a priori to be consistent with participant numbers used in previous sentence-picture verification (Clark & Chase, 1972) and eye-tracking studies (Papfragou et al., 2008).

**Stimuli and apparatus** Stimuli were presented using SR-Research Experiment Builder on a ViewSonic E70fB CRT monitor (1,024 x 768 pixels, 85-Hz refresh rate). The display subtended 23.8° x 17.2° at approximately 81 cm away from the monitor. Eye movements were recorded by an EyeLink 1000 Tower-Mounted eye tracker (1000-Hz sampling rate). The bars in the graph represented types of fruits (blueberries, limes, and oranges). The bars were the same width (1.5°) but differed in height; the shorter bar length was always 1.9°, and the taller bar length was always 5.5°. The bars were separated by 3.5° and appeared on unlabeled x- and y-axes.

**Procedure** The experiment began with a 9-point grid calibration sequence in which participants tracked a moving dot. Figure 1 shows a typical trial sequence. Participants first viewed a question about a magnitude relation (e.g., “Are there more blueberries than oranges?”). Questions were phrased equally often as *more* and *fewer* relations, and each fruit was assigned equally often as the relational target and reference object. Participants pressed a key when they were ready to view the graph. A fixation dot appeared at the center, and once participants pressed the spacebar and their eyes were positioned within 2° of fixation, a graph appeared. The depicted relation matched the question on half of the trials. The bar associated with the target appeared equally on the left and right sides. Participants were instructed to judge the relative heights of the bars (i.e., the taller bar represented a larger quantity). Participants responded “yes” or “no” by pressing one of two keys labeled “Y” and “N” as quickly but as accurately as possible. Based on the combination of fruit type, question quantifier, taller bar location, and correct response, there were 24 trials total divided into two blocks of 12 trials. Participants completed a second calibration sequence between the two blocks.

### Results

**Accuracy** Only trials in which saccades occurred (86 % of all trials) were included in all behavioral analyses. Mean accuracy on the task was 91.7 %.

**Eye movements** We only analyzed first saccades, excluding saccades that did not originate at central fixation (i.e., the start position was outside of a  $1.7^\circ$  window around central fixation) or that did not end on the display screen, which totaled 6 % of all first saccades. We examined saccades during graph viewing to test for biases. Saccade end positions were coded relative to central fixation ( $0^\circ$ ) in a binary fashion based on four bar properties: location (left = negative, right = positive), magnitude (short = negative, tall = positive), relational category (reference = negative, target = positive), and magnitude congruency, i.e., whether the bar's magnitude was incongruent (negative) or congruent (positive) with the question's quantifier (more/fewer). We tested for biases in the distribution of first saccades by comparing the average saccade size to  $0^\circ$ .

Average saccade biases are shown in Fig. 2a and b. Participants systematically looked first toward the relational target bar ( $M$  sentence-order bias =  $0.83^\circ$ ), 95 % confidence interval (CI) [ $0.57^\circ$   $1.08^\circ$ ],  $t(15) = 6.36$ ,  $p < 0.001$ , Cohen's  $d = 1.59$  (Fig. 2b). There were no systematic biases based on bar magnitude or magnitude congruency (both  $t$ 's  $< 0.5$ ) and only a mild tendency to look toward the left bar first ( $M$  spatial bias =  $-0.26^\circ$ ), 95 % CI [ $-0.6^\circ$   $0.08^\circ$ ],  $t(15) = 1.49$ ,  $p = 0.157$ , Cohen's  $d = 0.37$  (Fig. 2a). Thus, college-age participants tended to explore the graph using a visual routine that mimicked the color order of the question, regardless of whether the target bar appeared on the left or right side of the graph.

**Individual differences in visual routines** To explore the behavioral consequences of visual routines, we conducted two correlation analyses; one tested for an association between participants' average spatial bias and average response time (RT) and one tested for an association between average sentence-order bias and average RT (Fig. 2a and b). We excluded trials for which the RT was greater than 2 standard deviations above the mean RT for each participant and all incorrect trials; in total, 13 % of the dataset was excluded. There was a weak correlation between spatial bias and RT,  $r = -0.24$ ,  $t(14) = 0.93$ ,  $p > 0.25$ , whereas the correlation between sentence-order bias and RT was stronger,  $r = -0.49$ ,  $t(14) = 2.12$ ,  $p = 0.05$ , such that larger biases toward the sentence target bar were associated with faster responses. However, sentence-order biases were not significantly more predictive of average RTs than the spatial biases, Fisher's  $z = 0.84$ ,  $p > 0.25$ . Thus, although only the *identity* routine was related to participants' response efficiency, it was not a statistically stronger predictor of responses than the *spatial location* routine.

## Discussion

College-age viewers primarily extracted graph relations by inspecting bars using the sentence ordering of the bars' identity (color). Furthermore, individuals who used this *identity*

routine more systematically responded more quickly. The *identity* routine requires more attentional control, because the syntax of the question dictates which bar participants should attend to first and the viewer must inhibit a default routine to explore from left to right. Thus, individuals who have underdeveloped attentional control, such as young children (Diamond, Kirkham & Amso, 2002), may be less likely to use the *identity* routine to explore the graph. We repeated Experiment 1 with 6- and 8-year-olds to test whether children would be more likely to use default *spatial location* routines and whether larger spatial biases would be associated with slower responses.

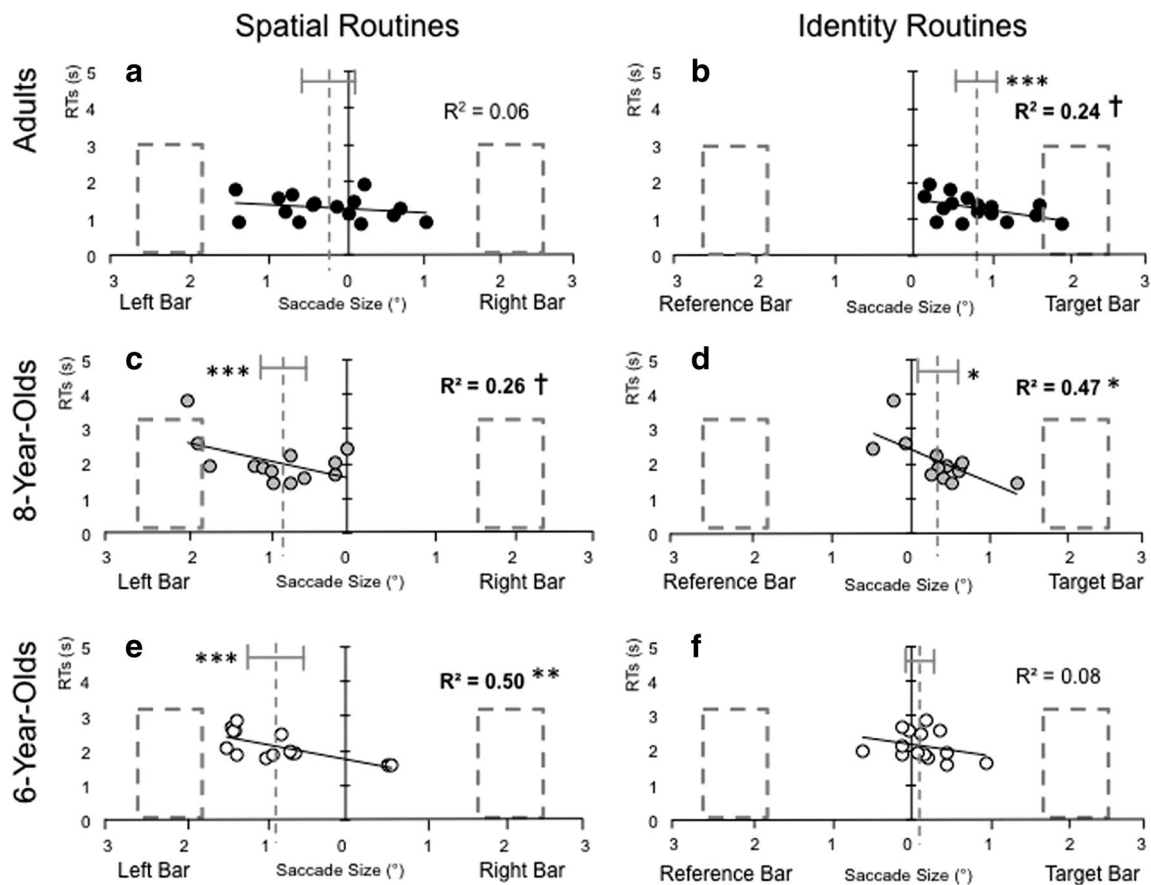
## Experiment 2

### Method

**Participants** We collected data from thirteen 8-year-olds (10 females, mean age = 102 months) and seventeen 6-year-olds (11 females, mean age = 78 months) who were recruited from the greater Chicago area by contacting their parents via email, direct mail, and pamphlets from the lab. All participants had normal or corrected-to-normal vision. Each child received a \$10 bookstore gift card as compensation. Data from four 6-year-olds were discarded; three children did not follow task instructions (one child performed at chance, and two children responded which fruit there was more of instead of responding "yes" or "no"), and one child did not maintain central fixation before viewing the graphs.

**Stimuli and apparatus** Participants were tested in a different laboratory room than adults to better accommodate the children. Stimuli were presented using SR-Research Experiment Builder on a ViewSonic E70fB CRT monitor (1,024 x 768 pixels, 85-Hz refresh rate), with the display subtending  $27^\circ \times 19.8^\circ$  at a distance of approximately 70 cm from the monitor. Eye movements were recorded by an EyeLink 1000 Desktop Mount eye tracker (1000-Hz sampling rate). The stimuli were identical to those in Experiment 1.

**Procedure** The procedures were identical to Experiment 1 with a few exceptions. The experimenter verified that participants understood the color-fruit mappings (e.g., blue bar = blueberries) and the magnitude mappings (e.g., taller bar = greater quantity). We modified the calibration so that it was framed as a game (see [Supplementary Materials](#) for details). This procedure was run before and after each block. The experimenter read each question out loud to ensure that participants understood the question. Once participants communicated that they were ready, the experimenter initiated presentation of a central fixation image, which consisted of several colored concentric circles (diameter =  $1.7^\circ$ ). Participants were



**Fig. 2** Eye movement results for Experiments 1 and 2. Dashed outlined bars represent graph stimuli at approximate locations of graph bars in the display and are not drawn to scale. The x-axes depict average saccade biases from 0 (initial center fixation) towards graph bars according to two potential eye movement patterns. (a), (c), and (e) present first saccades according to their horizontal spatial bias: first saccades directed left of center fixation plotted to the left versus those directed to the right of

fixation plotted to the right. (b), (d) and (f) present first saccades toward the relational “reference” bar plotted to the left versus those directed to the relational “target” bar plotted to the right. Each point shows the average saccade bias and average response time (y-axis) for a single participant. The vertical dotted line shows the average saccade size bias, collapsed across participants. Error bars are based on 95 % confidence interval

instructed to fixate this image as best they could. The graph appeared either once the child fixated the central image within 1° of center or once 2 seconds elapsed, whichever occurred first. After viewing the graph, the child responded to the question by saying “yes” or “no.” Verbal responses were recorded and timed with a stopwatch by a second experimenter who was unaware of the predicted outcomes of the experiment. To calculate RTs, the start time began once the graph appeared and ended once the child finished speaking.

**Results**

**Accuracy** Only trials in which eye movements occurred were included in all behavioral analyses. Mean accuracy was 96.7 % for 8-year-olds and 84.4 % for 6-year-olds. We conducted a one-way ANOVA on accuracy data with age group as a between-group factor (including adults, whose mean accuracy was 91.7 %). There was a significant effect of age group,  $F(2,40) = 4.96, p = 0.012, \eta_p^2 = 0.20$ ; post-hoc tests revealed

that 8-year-olds performed significantly better than 6-year-olds,  $t(25) = 2.93, p = 0.007, \text{Cohen's } d = 1.12$ , and marginally better than adults,  $t(27) = 1.89, p = 0.070, \text{Cohen's } d = 0.7$ , who trended towards higher performance than 6-year-olds,  $t(28) = 1.65, p = 0.111, \text{Cohen's } d = 0.6$ .

**Eye movements** We only included trials that contained unambiguously measurable saccades (i.e., the saccade start position occurred within a 1.7° window around the central fixation image), which occurred on 97 % of trials for 8-year-olds and on 79 % of trials for 6-year-olds.

We calculated the average size of first saccades based on various target properties in the same way as Experiment 1 to test for overall biases. In contrast to adults, 8-year-olds exhibited a significant left-bar-first bias ( $M$  spatial bias =  $-0.91^\circ$ ), 95 % CI [ $-1.25^\circ - 0.56^\circ$ ],  $t(12) = 5.07, p < 0.001, \text{Cohen's } d = 1.40$  (Fig. 2c). Eight-year-olds also showed a significant target-bar-first bias ( $M$  sentence-order bias =  $0.36^\circ$ ), 95 % CI

[0.12° 0.61°],  $t(12) = 2.89$ ,  $p = 0.014$ , Cohen's  $d = 0.80$  (Fig. 2d). No other biases were observed (all  $t$ 's < 1.3).

Similar to 8-year-olds, 6-year-olds showed a significant left-bar-first bias ( $M$  spatial bias =  $-0.86^\circ$ ), 95 % CI [ $-1.21^\circ$  0.50°],  $t(13) = 4.75$ ,  $p < 0.001$ , Cohen's  $d = 1.27$  (Fig. 2e) and a trend for a target-bar-first bias ( $M$  sentence-order bias =  $0.14^\circ$ ), 95 % CI [ $-0.05^\circ$  0.33°],  $t(13) = 1.46$ ,  $p = 0.168$ , Cohen's  $d = 0.55$  (Fig. 2f). No other biases were observed (all  $t$ 's < 1).

To test for developmental differences in overall biases, we conducted one-way ANOVAs on spatial and sentence-order biases with age group as a between-subjects factor (including data from Experiment 1). For spatial biases, there was a main effect of age group,  $F(2,40) = 4.24$ ,  $p = 0.021$ ,  $\eta_p^2 = 0.17$ ; adults had significantly weaker leftward biases than both 6-year-olds,  $t(28) = 2.36$ ,  $p = 0.026$ , Cohen's  $d = 0.86$ , and 8-year-olds,  $t(27) = 2.56$ ,  $p = 0.016$ , Cohen's  $d = 0.96$ , but spatial biases did not differ among 6- and 8-year-olds,  $t(25) = 0.21$ ,  $p > 0.25$ . Additionally, there was a main effect of age group on sentence-order biases,  $F(2,40) = 9.05$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.31$ , such that adults had significantly stronger sentence-order biases than both 6-year-olds,  $t(28) = 4.14$ ,  $p < 0.001$ , Cohen's  $d = 1.51$ , and 8-year-olds,  $t(27) = 2.54$ ,  $p = 0.017$ , Cohen's  $d = 0.95$ . Eight-year-olds had marginally stronger sentence-order biases than 6-year-olds,  $t(25) = 1.41$ ,  $p = 0.171$ , Cohen's  $d = 0.54$ .

Thus, viewing strategies for extracting graph relations changed over the course of development. On average, both 6- and 8-year-olds relied on the *spatial location* routine (attending to the left bar first), whereas adults showed a weaker left-bar-first bias. Additionally, both adults and 8-year-olds used the *identity* routine at rates above chance, but 6-year-olds did not. However, adults inspected the target bar first more systematically than 8-year-olds. Critically, these findings suggest that default visual routines diminish with age and are replaced with better routines that respect a task's context.

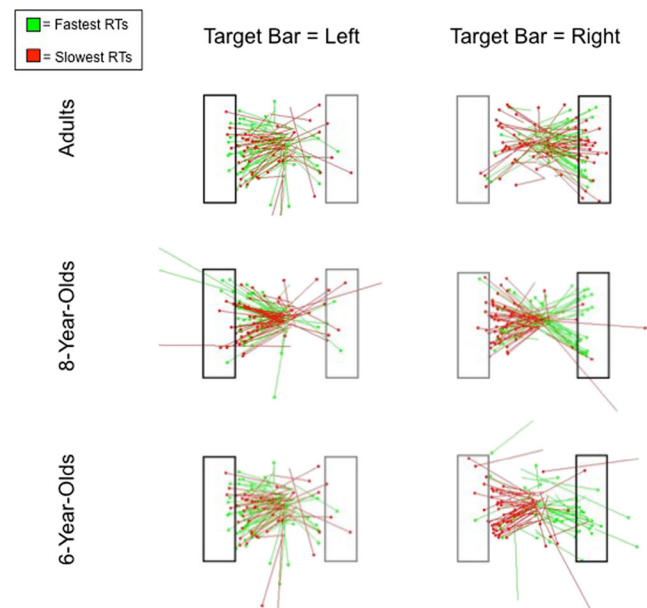
**Individual differences in visual routines** We correlated average RTs with spatial and sentence-order biases for 6- and 8-year-olds (Fig. 2c-f). We excluded RTs that were greater than 2 standard deviations above the mean RT for each participant and incorrect trials (totaling 12.7 % of all data). Spatial biases were marginally correlated with RTs for 8-year-olds,  $r = -0.51$ ,  $t(11) = 1.99$ ,  $p = 0.072$  (Fig. 2c) and significantly correlated with RTs for 6-year-olds,  $r = -0.71$ ,  $t(12) = 3.45$ ,  $p = 0.005$  (Fig. 2e). Eight-year-olds also showed a robust correlation between sentence-order bias and RT,  $r = -0.68$ ,  $t(11) = 3.10$ ,  $p = 0.010$  (Fig. 2d), whereas 6-year-olds did not,  $r = -0.29$ ,  $t(12) = 1.05$ ,  $p > 0.25$  (Fig. 2f). Although spatial and sentence-order bias correlations did not differ statistically for 8-year-olds, Fisher's  $z = 0.57$ ,  $p > 0.25$ , spatial bias was a significantly stronger predictor of RTs than sentence-order

bias for 6-year-olds, Fisher's  $z = 1.97$ ,  $p = 0.049$ . Thus, the more systematically children attended to the left bar first, the slower their response times were overall. In contrast, children who more successfully used the *identity* routine tended to be faster overall; however, only 8-year-olds showed this pattern.

**Effect of target bar location** We hypothesized that systematic left-bar-first routines would lead to slower responses particularly when the target bar appeared on the right side of the graph, because the bar colors would be encoded in the opposite order as the question. In support of this idea, we generally observed stronger correlations between sentence-order bias and average RT for displays in which the target bar appeared on the right than the left side (for details, see Figure S2 in the Supplemental Materials). However, the distribution of spatial biases differed between adults and children for target-right displays; whereas most adults tended to inspect the right bar first, most children showed strong left-bar-first biases.

To further probe how the target bar location affected RTs, we conducted a more fine-grained analysis of individual trials. For each age group, we divided RTs across participants into two categories based on speed of response: fastest (lowest third percentile) and slowest RTs (highest third percentile). We measured the distribution of fastest and slowest trials for which first saccades were launched toward the left versus right bar. Importantly, we performed this analysis separately for displays in which the target bar appeared on the left versus right.

Figure 3 plots first saccades color-coded by RT category (green = fastest, red = slowest) for each age group and display



**Fig. 3** Visualizations of first saccades by display type (target bar on the left or right side). Saccades are color coded by RT speed (red = slowest third of all RTs, green = fastest third of all RTs in each age group). Dots at the end of the lines represent the end points of the saccades



type. For target-right displays, children had more slow than fast RTs due to left-bar-first saccades. In contrast, adults had more fast than slow RTs when they inspected the right bar first. We verified these qualitative patterns with quantitative evidence by comparing the average number of fastest and slowest responses for saccades launched towards the left versus right bar as a function of target bar location (see Figure S3 in the Supplemental Materials available online). Thus, children incurred an RT cost, because they encoded the graph in the opposite order as the question, whereas adults showed an RT benefit, because they encoded the graph in the same order as the question. However, these RT effects only occurred when the target bar appeared on the right, suggesting that attentional control was more critical for these displays.

Although children showed stronger left-first biases than adults, it is possible that children intended to use the *identity* routine but were involuntary drawn toward the left side when the target bar appeared on the right side. If this were the case, then children may have spent a shorter amount of time fixating the left side before looking towards the right when the target bar appeared on the right side of the graph. To test this possibility, we conducted an ANOVA on the duration of leftward fixations with age as a between-group factor (6-, 8-year-olds, and adults) and target bar location (left or right) as a within-subjects factor. Only trials in which gaze shifted from the left to the right side of the graph were included, and if multiple consecutive leftward saccades occurred before the rightwards shift, fixation duration was summed across consecutive leftward saccades.

Based on these criteria, five adult participants were excluded from the analysis, because they never shifted their gaze from left to right in either or both target bar location conditions. There was a main effect of target location, such that leftward fixations were shorter when the target bar appeared on the right ( $M$  duration = 532 ms) than the left ( $M$  duration = 640 ms),  $F(1, 35) = 6.60, p = 0.015, \eta_p^2 = 0.16$ . There also was an interaction between age group and target location,  $F(2, 35) = 3.77, p = 0.033, \eta_p^2 = 0.18$ . Post-hoc tests revealed that only 8-year-olds had significantly shorter leftward fixations when the target bar appeared on the right ( $M = 474$  ms) versus left ( $M = 718$  ms),  $t(12) = 3.98, p = 0.002$ , Cohen's  $d = 1.35$ , although 6-year-olds showed a trend for shorter leftward fixations for target-right ( $M = 526$  ms) than target-left displays ( $M = 621$  ms),  $t(13) = 1.56, p = 0.142$ , Cohen's  $d = 0.51$ . In contrast, for adults, leftward fixation durations were statistically similar for target-right ( $M = 606$  ms) and target-left displays ( $M = 564$  ms),  $t(10) = 0.43, p > 0.25$ ; however, because adults rarely attended to the left bar first for target-right displays, it is possible that we did not observe any differences in left bar gaze duration because of low power. Because children were more likely to fixate the left bar for a shorter duration when the target appeared on the right than left, it is possible that children involuntarily shifted attention towards the left bar

and quickly reoriented attention towards the right bar. Additionally, 8-year-olds showed a stronger reorienting effect than 6-year-olds, suggesting that the ability to compensate for default routines also may improve with age.

## General discussion

Relational perception is a severely capacity limited process. Yet, dozens of relational descriptions can be pulled from even the simplest displays (Taylor & Tversky, 1992). We show that effectively extracting the correct relation for a task relies on the way in which a viewer moves their attention through the values. We asked college students and elementary school-aged children to verify that a graph agreed with a simple statement, e.g., "Are there more blueberries than oranges?" and found that moving attention in the "correct" way—shifting one's eyes to the relation's target—predicted fast response times. The "incorrect" strategy—shifting one's eyes to the left bar of the graph, regardless of the question asked—led to slow response times; this pattern was especially prevalent among 6-year-olds. Strategy choice had larger behavioral consequences when the target bar was on the right side; whereas children incurred an RT cost by attending to the left (mismatched) bar first, adults experienced an RT benefit by attending to the right (matched) bar first. Children also were more likely to compensate for initial leftward saccades by quickly shifting to the right bar for target-right displays; however, 8-year-olds showed stronger reorienting effects than 6-year-olds. Together, we show that it is easier to coordinate visual relations with verbal information when the two representations are encoded in the same serial order. However, this temporal alignment requires attentional control, which is limited in younger children.

Our findings have several implications for the practice of integrating relations from multiple representations. For instance, coordinating relations in text with visual relations may be easier if people read text first so that they can attend strategically to visualizations (Bergey, Cromley & Newcombe, 2015; Bacharach, Carr & Mehner, 1976). Efficient extraction of relations may be especially critical for processing more complex displays or for generating inferences about a visualization. Additionally, educators may benefit from knowing the potentially negative consequences of using default left-first routines when teaching young children how to interpret visualizations. Child participants may have relied more on automatic spatial routines, because children often scan from left to right when learning to write and read text, charts, and tables in English-speaking cultures (Tversky, Kugelmass & Winter, 1991). It therefore is possible that preliterate children might exhibit weaker spatial biases when viewing graphs; on the other hand, preliterate English-speaking children still demonstrate leftward biases in tasks

other than reading and writing, such as object counting (for review, see Nuerk et al., 2015). Thus, further research is necessary to determine whether children younger than age 6 would show left-first biases when viewing graphs. Another open question is how strongly children would rely on vertical spatial routines, such as attending to a top bar first if the bars extended horizontally (a “row graph” as opposed to a “column graph”). Previous studies have established an attentional bias for the top in adults (Clark & Chase, 1972; Xu & Franconeri, 2015). Children also might show an upwards bias, because many school-related materials, such as text and charts, begin at the top (Tversky, 2011). Because there is a strong mapping between vertical space and quantity (Clark & Chase, 1972; Tversky, 2011), it is also possible that children might rely on the quantifier of the question to guide eye movements for vertically arranged bar graphs (e.g., showing a top bias for *more* questions and a bottom bias for *fewer* questions). Although it may be possible to teach children to guide their attention in a more efficient manner during graph comprehension, another solution is to design displays containing text and visualizations for younger children such that a spatial routine can be effective in comparing the relations expressed in each format.

**Acknowledgments** This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant IES-R205A120531 to Northwestern University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors thank Karly Raber, Sherry Vernon, and Ernesto Melchor for data collection assistance and Lei Yuan and Linsey Smith for helpful comments.

## References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183–198.
- Bacharach, V. R., Carr, T. H., & Mehner, D. S. (1976). Interactive and independent contributions of verbal descriptions to children's picture memory. *Journal of Experimental Child Psychology, 22*(3), 492–498.
- Bergey, B. W., Cromley, J. G., & Newcombe, N. S. (2015). Teaching high school biology students to coordinate text and diagrams: Relations with transfer, effort, and spatial skill. *International Journal of Science Education, 37*(15), 2476–2502.
- Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*(2), 75–100.
- Clark, H., & Chase, W. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 5*(17), 472–517.
- 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.
- Dickinson, C. A., & Intraub, H. (2009). Spatial asymmetries in viewing and remembering scenes: Consequences of an attentional bias? *Attention, Perception, & Psychophysics, 71*, 1251–1262.
- Feeney, A., & Webber, L. J. (2003). Analogical representation and graph comprehension. *Lecture Notes in Computer Science, 2733*, 212–221.
- Franconeri, S. L., Scimeca, J. M., Roth, J. C., Helseth, S. A., & Kahn, L. E. (2012). Flexible visual processing of spatial relationships. *Cognition, 122*, 210–227.
- Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124–158.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155–170.
- Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language, 57*(4), 544–569.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274–279.
- Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14*(11), 497–505.
- Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*(3), 480–517.
- Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Cognitive Science, 14*, 153–157.
- Kuhn, D., & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of Cognition and Development, 1*, 113–129.
- Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1015–1036.
- Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28*(2), 103–174.
- Luck, S. J., & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Science, 95*, 825–830.
- Matthews, W. J., & Dylman, A. S. (2014). The language of magnitude comparison. *Journal of Experimental Psychology: General, 143*(2), 510–520.
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science, 21*(2), 71–77.
- Nuerk, H.-C., Patro, K., Cress, U., Schild, U., Friedrich, C. K., & Göbel, S. M. (2015). How space-number associations may be created in preliterate children: Six distinct mechanisms. *Frontiers in Psychology, 6*(215), 1–6.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184.
- Roth, J. C., & Franconeri, S. L. (2012). Asymmetric coding of categorical spatial relations in both language and vision. *Frontiers in Psychology, 3*(464), 1–14.
- Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition, 20*(5), 483–496.
- Townsend, D. J. (1974). Children's comprehension of comparative forms. *Journal of Experimental Child Psychology, 18*(2), 293–303.
- Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science, 3*, 499–535.
- Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graph productions. *Cognitive Psychology, 23*, 515–557.
- Ullman, S. (1984). Visual routines. *Cognition, 18*, 97–159.
- Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics, 63*(3), 381–389.
- Xu, Y. Q., & Franconeri, S. L. (2015). The capacity for visual features in mental rotation. *Psychological Science, 26*(8), 1241–1251.