Matlen, B. J., Gentner, D., & Franconeri, S. L. (2020). Spatial alignment facilitates visual comparison. *Journal of Experimental Psychology: Human Perception and Performance,* 46(5), 443-457.

American Psychological Association

Journal of Experimental Psychology: Human Perception and Performance

ISSN: 0096-1523

2020, Vol. 46, No. 5, 443–457
http://dx.doi.org/10.1037/xhp0000726

Spatial Alignment Facilitates Visual Comparison

Bryan J. Matlen
WestEd, Redwood City, California

Dedre Gentner and Steven L. Franconeri
Northwestern University

Humans have a uniquely sophisticated ability to see past superficial features and to understand the relational structure of the world around us. This ability often requires that we compare structures, finding commonalities and differences across visual depictions that are arranged in space, such as maps, graphs, or diagrams. Although such visual comparison of relational structures is ubiquitous in classrooms, textbooks, and news media, surprisingly little is known about how to facilitate this process. Here we suggest a new principle of spatial alignment, whereby visual comparison is substantially more efficient when visuals are placed perpendicular to their structural axes, such that the matching components of the visuals are in direct alignment. In four experiments, this direct alignment led to faster and more accurate comparison than other placements of the same patterns. We discuss the spatial alignment principle in connection to broader work on relational comparison and describe its implications for design and instruction.

Public Significance Statement

This research reveals that the way in which visuals are placed on the page or screen influences the efficiency by which people can identify similarities and differences between the visuals. Arranging visuals such that their corresponding components can be readily aligned optimizes the efficiency of visual comparison.

Keywords: spatial analogy, comparison, structure-mapping, visualization

Supplemental materials: http://dx.doi.org/10.1037/xhp0000726.supp

Humans have a uniquely sophisticated ability to see past superficial features and to understand the relational structure of the world around us. Reasoning about relational structure is often most powerful when supported by visuospatial representations, such as maps, graphs, and diagrams (Ainsworth, Prain, & Tytler, 2011; Gattis, 2002; Hegarty & Just, 1993; Kellman, Massey, & Son, 2010; Tversky, 2011; Uttal, Fisher, & Taylor, 2006).

A particularly powerful way to gain insight into relational structure is to compare visual representations (Alfieri, Nokes-Malach, & Schunn, 2013; Gentner & Namy, 1999; Kurtz & Gentner, 2013; Rau, 2017). For example, people who compare simultaneously presented visual examples of heat-flow are more likely to notice the common phenomena than are those who describe the visuals separately, suggesting that comparison highlights commonalities (Kurtz, Miao, & Gentner, 2001). Visual comparison can also highlight differences (Gentner et al., 2016; Sagi, Gentner, & Lovett, 2012). For example, medical students who compared X-rays of diseased lungs with those of healthy lungs were subsequently better able to identify focal lung diseases in further X-rays (Kok, de Bruin, Robben, & van Merriënboer, 2013). Though humans can detect changes in sequentially presented visuals, comparison of simultaneously presented visuals allows learners to more fluently encode and compare images (e.g., Christie & Gentner, 2010; Larsen, McIlhagga, & Bundesen, 1999). Simultaneous presentation likely minimizes the impact of working memory limitations that impede comparison across sequentially presented displays (Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009). An advantage of simultaneous visual presentation over sequential presentation for learning has been found for a variety of educational domains (e.g., Brooks, Norman, & Allen, 1991; Gadgil,

Bryan J. Matlen, STEM Program, WestEd, Redwood City, California; Dedre Gentner and Steven L. Franconeri, Department of Psychology, Northwestern University.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A170165 to WestEd. The opinions expressed are those of the authors and do not necessarily represent views of the Institute or the U.S. Department of Education. The research was also supported by National Science Foundation (NSF) Grant SBE-0541957 the Spatial Intelligence and Learning Center (SILC), by ONR Grant N00014-16–1-2613 to Dedre Gentner, and by NSF Grants CHS-1901485 and NSF EHR DRL-1661264 to Steven L. Franconeri. We thank Kevin Huang and Larry Thomas for statistical consultation, and Nina Simms for assistance with data collection.

Data and stimuli used in this research are available through the Center for Open Science https://osf.io/sk9vt (Matlen, 2019).

Correspondence concerning this article should be addressed to Bryan J. Matlen, STEM Program, WestEd, 400 Seaport Court, Suite 222, Redwood City, CA 94063. E-mail: bmatlen@wested.org

Nokes-Malach, & Chi, 2012; Jee et al., 2014; Kellman, 2013; Matlen, Vosniadou, Jee, & Ptouchkina, 2011; Rittle-Johnson & Star, 2007).

Given that simultaneous visual comparison is a critical format for learning, it is important to investigate the most effective ways to present such comparisons. In visual figures, such as those shown in Figure 1, much of the critical information is conveyed by the spatial configuration: For example, the molecular notation of Ö = C = Ö differs importantly from Ö = Ö = C. Comparing two visuals requires comparing not only their concrete elements, but also the spatial relational structure of those elements. As another illustration, in the first column of Figure 1, comparing the three colored bars (dark gray, black, and light gray) to the three legend categories (A, B, and C) is easiest when the viewer can quickly match each of the elements in one set to their corresponding elements in the other set, allowing the viewer to carry out only three comparisons, out of a total of nine possible pairings of elements between the graphed values and legend entries (three appropriate, and six inappropriate). But if this matching process is not efficient, the viewer could be slowed by taking the time to compare inappropriate pairings. This inefficient matching may also increase the likelihood of making erroneous matchings. Thus, visual designs that facilitate efficient spatial matching should speed up comparison in displays that present visuals simultaneously.

This process has been studied as a process of analogical comparison—an alignment of common relational structure (Gattis, 2002; Yuan, Uttal, & Gentner, 2017). During comparisons, people implicitly seek a one-to-one mapping in which like relations are put into correspondence (Doumas, Hummel, & Sandhofer, 2008; Gentner, 1983, 2010; Gentner & Markman, 1997; Jones & Love, 2007; Kokinov & French, 2003; Krawczyk, Holyoak, & Hummel, 2005; Markman & Gentner, 1993; Sagi et al., 2012; Thibaut, French, & Vezneva, 2010). In a visual comparison, this mapping should be most fluent when visual representations are placed so that their corresponding elements and relations are readily matched. That is, the comparison process should be more fluent to the degree that (1) the intended relational correspondences are readily apparent and (2) potential competing correspondences are minimized.

This leads to our central claim. We propose the *spatial alignment principle*—that visual comparison is more efficient when the visual representations have their principle axes parallel and are placed orthogonally to their principle axes (see Figure 1). For example, in the graph comparison, the viewer should match the dark gray bar with the A legend entry, the black bar with the B legend entry, and so on. As shown in the first row of Figure 1, the axis along which information changes is horizontal. Thus, according to the spatial alignment principle, the fluency of this spatial alignment process should be greatest when the visuals are placed vertically with their axes parallel, which we call a *direct alignment*. This design allows corresponding elements to be easily matched by their identical horizontal position in the display. Likewise, if the structural axes were vertical (e.g., if the examples were rotated 90 degrees), then the optimal placement would be horizontal—placed side by side (second row of Figure 1), allowing matches by each corresponding element's vertical position in the display.

The spatial alignment hypothesis holds that visual comparison is most efficient when the two visuals are in direct alignment, and least efficient when the visuals are in impeded alignment. In direct spatial alignment, both fluency criteria are satisfied: (1) the corresponding elements and relations are juxtaposed, and (2) corresponding elements are relatively far from other similar but noncorresponding elements that might compete with the intended mapping. Impeded alignment should be difficult, because matching horizontal or vertical positions are no longer diagnostic for matching corresponding elements, and matches should compete

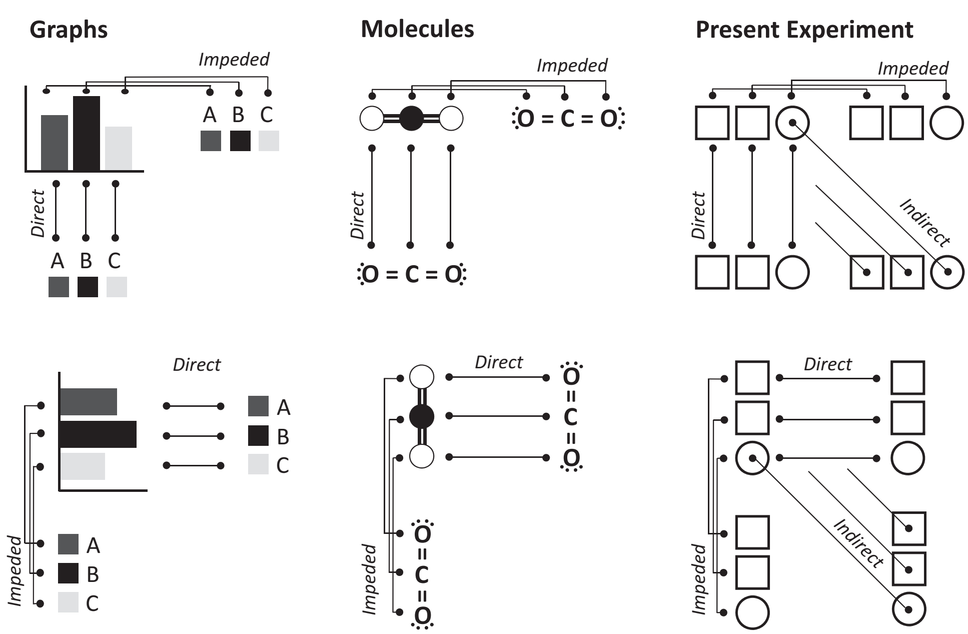

Figure 1. Direct and impeded placements for visuals with horizontal axes (top row) and vertical axes (bottom row), consisting of a bar graph and its legend entries (left), molecular notations (middle), and shapes used in the present set of experiments (right).

for attention with the intervening potential correspondences. Our experiments also test an intermediate case, indirect alignment. In indirect alignment, as in the rightmost example of Figure 1, the oblique angle between the two visuals makes it more difficult to differentiate competing correspondences.

Though some research has explored how spatial alignment of visuals affects comparison (e.g., Hribar, Haun, & Call, 2012; Larsen & Bundesen, 1998; Paik & Mix, 2008), we know of no work that systematically manipulates both the axis and placement of visuals to examine the impact of spatial alignment on visual comparison. To our knowledge, the current research is the first systematic investigation of spatial alignment.

We test three predictions of the spatial alignment principle across four studies, using a same-different task over sets of simple shapes or colors. In Experiment 1, we test the prediction that comparison should be more efficient for direct alignment than for impeded alignment, and possibly also more efficient than indirect (oblique) alignment. Experiment 2 tests whether the impeded condition is less efficient specifically due to competing correspondences, as we predict, or instead because it places irrelevant barriers in the way of relevant comparisons. Experiment 3 tests another prediction of the spatial alignment hypothesis: namely, that as in other instances of analogical mapping, it should apply to purely relational comparisons that lack any concrete matches. In Experiment 4, we directly compare spatial alignment effects between comparisons with object matches and comparisons that are purely relational, in a within-subjects manipulation. Experiment 4 also provides an opportunity to replicate the findings of Experiments 1 and 3.

Experiment 1

In Experiment 1 we test the spatial alignment hypothesis by varying how visual comparisons are placed. Subjects make same–different judgments (Farell, 1985) about pairs of triplets composed of basic shapes or colors. We manipulate the structural axes of the triplets and their spatial placement. We predict more efficient comparison, in terms of speed, accuracy or both, for vertically arranged triplets with horizontal structural axes, and for horizontally arranged triplets with vertical structural axes.

Method

Participants. Participants were 16 adults (M = 19.3 years, range = 18–23 years, eight women, eight men) from Northwestern University, who each received course credit or payment. Two participants were unable to finish due to technical malfunction.

Based on this resulting sample size and the experimental design, we conducted a sensitivity analysis for the key prediction (i.e., the placement x triplet interaction, described below). We used G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) to conduct an F test for repeated-measures analysis of variance (ANOVA), specifying a within-between interaction, with two groups (triplets) and four measurements (placements). Assuming error probability of .05, nonsphericity correction of 1, and power of .80, the Experiment was 80% powered to detect an effect of Cohen's f = .33.

Materials and design. All stimuli and trials were created using MATLAB. On each trial, a pair of triplets was presented for a same–different judgment. Half the trials were shape trials and half were color trials (referred to as *stimulus type*). Each trial consisted of a pair of triplets that varied in (1) the structural axes of the triplets (vertical or horizontal)—referred to as *triplet*; (2) pair placement (horizontal, vertical, or oblique)—referred to as *placement*; and (3) whether the two triplets were the same or different—referred to as *concordance*. Throughout these studies, within each pair, the triplet axes were parallel (either both vertical or both horizontal). The design was 4 Placement (horizontal, oblique 1, oblique 2, vertical) × 2 Triplet (horizontal vs. vertical) × 2 Stimulus Type (colors vs. shapes) × 2 Concordance (different vs. same), all within-subjects. Each placement condition comprised a fourth of the total trials, and all participants completed each of these four conditions for a total of 864 trials.

Triplets. In shape trials, each triplet consisted of three black geometric shapes (triangles, squares, or circles). In color trials, each triplet consisted of star shapes varying in color (red [RGB: 228, 26, 28], blue [RGB: 55, 126, 184], or green [RGB: 77, 154, 174]). For simplicity, we detail the makeup of shape trials (see Figure 2), but the same plan was used for the color trials. Each shape triplet was made up of two geometric shapes—two alike and one different, in a particular order (triangle–triangle–square, triangle–square–triangle, or square–triangle–triangle). Thus, there were 18 possible triplets (six possible pairings of two out of three geometric shapes, and three orderings within each of these).

Pairs of triplets always shared the same principle axes—horizontal or vertical—and contained the same shapes; however, the order of shapes could vary between the triplets (in which case, a "different" response was required). By varying the placement and triplet axes, we created three types of spatial alignments: direct, indirect, and impeded. In direct trials, the placement of triplets was perpendicular to their axes (e.g., horizontal triplets placed vertically, and vice versa). In impeded trials, the placement of triplets was parallel to their axes (e.g., horizontal triplets placed horizontally and vice versa). In indirect trials, the placement was oblique to the axes of the triplets.

Triplets were displayed at 1,024 × 768 resolution on a 17-in. monitor. Displays consisted of two triplets displayed on a white background, centered around a black fixation point (16 pixels × 16 pixels in width and height). Within triplets, objects had 10 pixels of spacing between their 74-pixel × 74-pixel bounding boxes (either horizontally or vertically). Thus, vertical triplets measured 74 pixels × 242 pixels, and horizontal triplets measured 242 pixels × 74 pixels. Triplet centers could be 205 pixels from fixation and the distances between triplet centers were 410 pixels across all trials.

Procedure. Participants were tested in quiet room at Northwestern University. The experiment was delivered on a computer running E-Prime. Participants were told that they were going to be making simple same and different judgments of shape and color sequences—they were instructed to press *1* if the sequences were the same and *0* if they were different. Participants were asked to respond as quickly and accurately as possible. After reading the instructions, participants completed six practice trials, using triplets with different elements than those in the actual experiment. Shape practice trials consisted of octagons and stars, and color practice trials consisted of black and white stars. Participants received feedback on the speed and accuracy of their responses after each practice trial.

After completing the practice trials, participants were told that they would begin the experiment, and were asked if they had any

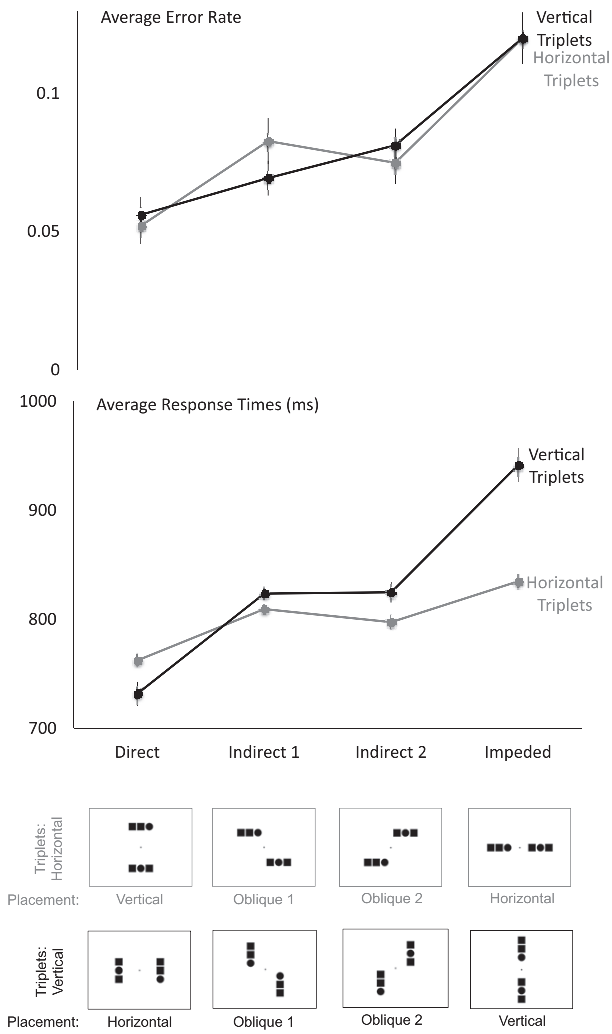

Figure 2. Results of Experiment 1 across spatial placements for the two triplet conditions: average proportion of error (top panel) and average response time (in ms; bottom panel). Examples of each triplet and placement condition are shown at bottom. Error bars represent within-subject standard errors (Cousineau, 2005).

questions. Participants were told that there was no time limit on their responses and were reminded to respond as quickly and accurately as possible. They were encouraged to take short breaks between trials if they found themselves becoming inattentive or drowsy. Each trial consisted of a fixation cross presented for 1,000 ms. The stimuli for the trial appeared immediately after fixation, at which point participants made a timed same–different decision. After each response, participants were reminded about which button to press for same and different trials and were asked to press the space bar to begin the next trial. No feedback was given during experimental trials.

Stimulus type (i.e., shape or color) was blocked and order was counterbalanced across participants: Half the participants received shape and then color trials; the other half received the reverse order. All other conditions were presented randomly within the blocks. After completing the first block (e.g., shapes), participants were told that they would again complete the same task, but with triplets that varied along the other stimulus dimension (e.g., color). As in the first block, participants completed six practice trials of the remaining stimulus type (with feedback) and then received the experimental trials (without feedback).

Results

Data for all experiments were analyzed after all participants had completed the study using R statistical software Version 3.5.2 (R Core Team, 2018). Independent sample *t* tests of the effects of block order were not significant for either response time or accuracy ($ts < 1.18$, $ps > .24$); therefore, this factor was dropped from further analyses. We first present response time results and then accuracy.

Response time. We eliminated any trial with response time at or above 5,000 ms (<1% of trials). On average, correct response times were 816 ms ($SD = 263$ ms). Response times for correct trials within each stimulus condition were averaged within each subject. We then conducted a 2 (stimulus type) × 4 (placement) × 2 (triplet) × 2 (concordance) repeated-measures ANOVA. This analysis revealed two main effects: Responses were faster for horizontal triplets ($M = 801$ ms, $SD = 135$ ms) than vertical triplets ($M = 831$ ms, $SD = 165$ ms) as confirmed by a main effect of triplet, $F(1, 13) = 16.93$, $p = .001$, $f = .95$, and speed of responses differed between placement conditions (vertical $M = 852$ ms, $SD = 183$ms; oblique 1 $M = 817$ ms, $SD = 135$ ms; oblique 2 $M = 811$ ms, $SD = 131$ ms; horizontal $M = 783$ ms, $SD = 142$ ms) as confirmed by a main effect of placement, $F(3, 39) = 21.79$, $p < .001$, $f = 1.58$. There were three significant interactions: Triplet × Stimulus Type, $F(1, 13) = 19.85$, $p = .001$, $f = .49$, Placement × Concordance, $F(3, 39) = 6.97$, $p = .001$, $f = .74$, and the predicted Triplet × Placement interaction, $F(3, 39) = 50.48$, $p < .001$, $f = 3.25$ (see Figure 2).

Because we predicted the Triplet × Placement interaction, we explore this interaction in depth (other significant interactions are described in the online supplementary materials). Based on our hypotheses regarding spatial alignment, we grouped the conditions according to their spatial alignment conditions. Horizontal triplets in vertical placements (and vice versa) were categorized as direct trials, horizontal triplets in horizontal placements (or vertical in vertical) were categorized as impeded trials, and triplets that were in either of the oblique placements were categorized as indirect trials. We then compared response times for the spatial alignment conditions within each triplet condition. For horizontal triplets, participants were faster for direct ($M = 762$ ms, $SD = 118$ ms) than for both indirect trials ($M = 804$ ms, $SD = 113$ ms) and impeded trials ($M = 835$ ms, $SD = 132$ ms; $ts > 5.45$, $ps < .001$), for indirect versus direct ($d = .35$, 95% CI [.21, .48]), for impeded versus direct ($d = .55$, 95% CI [.37, .73]). This pattern was the same for vertical triplets: Participants were faster for direct ($M = 732$ ms, $SD = 103$ ms) relative to both indirect ($M = 824$ ms, $SD = 126$ ms) and impeded trials ($M = 942$ ms, $SD = 169$ ms; $ts > 9.62$, $ps < .001$), for indirect versus direct ($d = .63$, 95% CI [.48, .77]), for impeded versus direct ($d = 1.18$, 95% CI [.79, 1.57]). For both horizontal and vertical triplets, participants were also faster for indirect relative to impeded conditions ($ts > 3.48$, $ps < .005$), for horizontal ($d = .21$, 95% CI [.09, .34]), for vertical ($d = .63$, 95% CI [.41, .84]).

Accuracy. To explore accuracy, proportion correct within each stimulus condition were averaged within each subject. We then conducted a 2 (stimulus type) × 4 (placement) × 2 (triplet) × 2 (concordance) repeated-measures ANOVA. This analysis revealed three significant interactions: a Triplet × Stimulus Type interaction, $F(1, 13) = 7.93$, $p = .02$, $f = .38$, qualified by a Triplet × Stimulus × Concordance interaction, $F(1, 13) = 9.77$, $p = .008$, $f = .53$ (both interactions are described in more detail in the online supplementary materials) and the predicted Triplet × Placement interaction, $F(3, 39) = 16.51$, $p < .001$, $f = 1.22$.

To explore the predicted Triplet × Placement interaction, we again grouped conditions based on their spatial placement (i.e., direct, indirect, and impeded). For horizontal triplets, participants made fewer errors on direct ($M = .05$, $SD = .05$) than on indirect ($M = .08$, $SD = .06$) and impeded trials ($M = .12$, $SD = .07$, $ts > 2.83$, $ps < .05$), for indirect versus direct ($d = .49$, 95% CI [.11, .87]), for impeded versus direct ($d = 1.09$, 95% CI [.56, 1.62]). This pattern was the same for vertical triplets: Participants made marginally fewer errors on direct ($M = .06$, $SD = .05$) than on indirect trials ($M = .08$, $SD = .05$), and fewer errors on direct relative to impeded trials ($M = .12$, $SD = .07$, $ts > 2.14$, $ps \leq .05$), for direct versus indirect ($d = .40$, 95% CI [.00, .79]), for direct versus impeded ($d = .94$, 95% CI [.46, 1.42]). For both horizontal and vertical triplets, participants made fewer errors on indirect relative to impeded trials ($ts > 3.22$, $p < .01$), for horizontal ($d = .65$, 95% CI [.19, .1.11]), for vertical ($d = .65$, 95% CI [.22, 1.08]).

Summary. Consistent with the spatial alignment principle, participants were faster and more accurate for direct spatial alignment than for either indirect or impeded alignment. They were also faster and more accurate for indirect relative to impeded trials.

Experiment 2

The results of Experiment 1 are consistent with our prediction that direct spatial alignment aids analogical comparison, and impeded spatial alignment interferes. However, a potential alternative explanation is that the adverse effects in the impeded condition are due to simple visual blocking, rather than to difficulty aligning the two triplets. For example, a person comparing the impeded pair ABA–ABB must note that the final elements of the two triplets do not match in order to correctly respond "different." Perhaps the presence of intervening items renders this more difficult. Experiment 2 tested these possibilities by comparing impeded pairs with pairs in which visual barriers have been placed between the two triplets. Further, we explore whether the barriers or competing correspondences must be spatially between the two members of the pair in order to interfere, versus simply being physically near them.

Method

Participants. Participants were 22 adults ($M = 24.47$ years, range = 18 – 48 years, 16 women, six men) recruited from Northwestern University. Based on this sample size and the experimental design, we conducted a sensitivity analysis for the main effect of condition (described subsequently). We used G*Power 3.1 to conduct an F test for a repeated-measures ANOVA, specifying within-factors, one group, and five measurements (conditions). Assuming error probability of .05, nonsphericity correction of 1, and power of .80, the Experiment was 80% powered to detect an effect of Cohen's $f = .24$.

Conditions. Because there were no main effects of stimulus type in Experiment 1, we used only shape trials in Experiment 2. The impeded condition was as in Experiment 1. There were four other conditions—all in direct alignment—to which a third element was added. To test whether the impedance effect in Experiment 1 was due to the presence of competing potential correspondences, in two of these conditions an additional triplet like the ones being compared (pattern condition) was added; in the other two, a solid rectangle was added (solid condition; Figure 3). If, as predicted, the impedance effect results from competing potential correspondences, then the triplet, but not the barrier, should lead to interference. To test whether the impedance effect resulted specifically from intervening elements, the additional element was placed either between the compared triplets (barrier condition), or to the side (nonbarrier condition). These factors were crossed to create four conditions: solid-nonbarrier, solid-barrier, pattern-nonbarrier, and pattern-barrier, in addition to an impeded condition with only two triplets (see Figure 3). Thus, the design was five

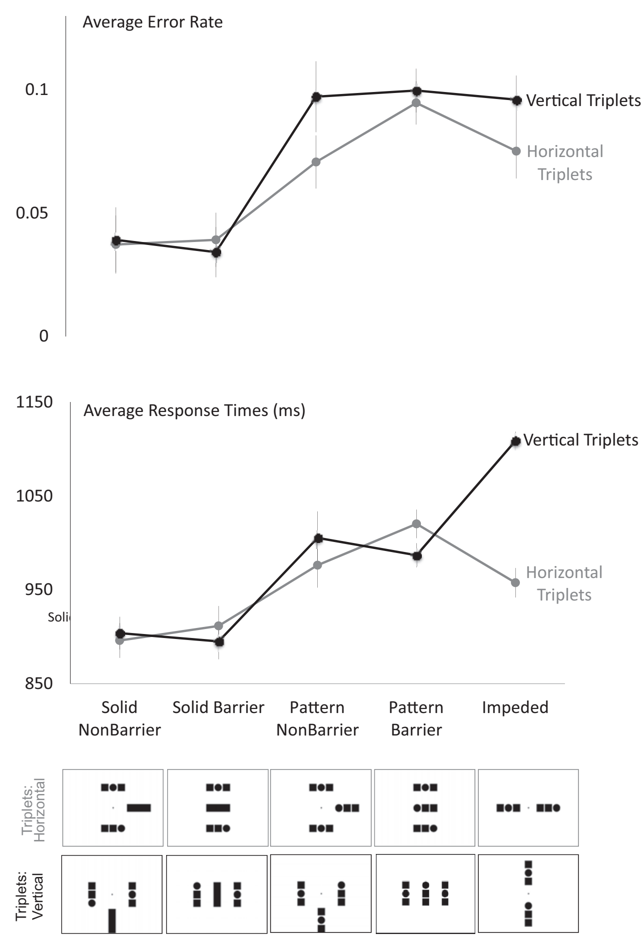

Figure 3. Results of Experiment 2 across conditions for the two triplet types. The top panel displays error rate and the bottom panel displays average response time (in ms). Examples of the of each triplet and condition type are shown at bottom. Error bars represent within-subject standard errors.

Condition × 2 Triplet (horizontal or vertical) × 2 Concordance (same or different), all within-subjects.

Materials and procedure. The design of triplets was as in Experiment 1. For every shape combination (e.g., triangle–triangle–square), there were six different trials and three same trials. Unlike Experiment 1, trials in Experiment 2 consisted only of direct and impeded spatial placements (i.e., there were no indirect placements). In addition, direct trials always contained a third element (the "distracting element"), which could be either solid or in a triplet pattern.

Distracting elements that were solid consisted of a black, rectangular block measuring 74 pixels × 242 pixels for vertically oriented stimuli (the same size as the compared triplets) and these dimensions were reversed for horizontally oriented stimuli. Distracting elements that were patterns were also 74 pixels × 242 pixels but were triads of the same shapes used in compared triplets. For example, if the compared triplets consisted of two triangles and a square, the patterned distracting element also consisted of two triangles and a square.

Distracting elements also varied in whether they were barriers or nonbarriers. Barriers were distracting were placed between the compared triplets, such that the direct correspondence lines physically crossed through the barrier. Nonbarriers were placed outside of the correspondence lines of the compared triplets by 10 pixels (counterbalanced to be above or below the compared triplets for vertical triplets and left or right of the compared triplets for horizontal triplets). Impeded trials were identical to those of Experiment 1 and contained no distracting elements. Thus, Experiment 2 consisted of 540 trials (5 Conditions × 6 Shape Sets × 9 Trials × 2 Triplet Axes) presented in a random order.

The procedure for Experiment 2 was identical to that of Experiment 1 except that participants received more extensive instruction before the experimental trials. We believed this added instruction was necessary as the introduction of a distracting element increased the complexity of the task. As in Experiment 1, participants were told that they would make same or different judgments of triplets that consisted of basic shapes and were given examples of same and different trials. They were then told that there would be a distracting element on some trials, but that their task remained the same: to determine whether the outside triplets (referred to as the *key images* in the instruction) were the same or different. Diagrams of example trials were shown to participants with both key triplets and distracting elements labeled, and correct responses were provided. After completing the instruction phase (approximately 5 min), participants were given 15 practice trials (three for each condition) with accuracy and response time feedback.

Results

To preview, we found that the presence of competing patterns was more detrimental to performance than the presence of solid blocks. Thus, the process is sensitive to competing correspondences, as predicted by the spatial alignment hypothesis. It did not matter whether these competing patterns were placed between the two triplets, or off to the side.

Response time. We first analyzed response time for correct trials by averaging response times within each condition and subject and conducting a 2 (Triplet) × 2 (Concordance) × 5 (Condition) within-subjects ANOVA. This analysis revealed main effects of triplet and condition (see Figure 3). Responses were faster for horizontal triplets (M = 952 ms, SD = 265 ms) than for vertical triplets (M = 980 ms, SD = 305 ms) as confirmed by a main effect of triplet, $F(1, 21) = 8.48$, $p = .008$, $f = .42$, and speed of responses differed between conditions, $F(4, 84) = 10.30$, $p < .001$, $f = 1.65$, impeded (M = 1,033 ms, SD = 262 ms), pattern barrier (M = 1,004 ms, SD = 306 ms), pattern nonbarrier (M = 991 ms, SD = 348 ms), solid barrier (M = 903 ms, SD = 241 ms), solid nonbarrier (M = 900 ms, SD = 237 ms). These main effects were qualified by a significant interaction between triplet and condition, $F(4, 84) = 15.41$, $p < .001$, $f = .99$.

To explore the interaction, we conducted two further analyses. First, we compared each distracting element condition to the impeded condition. Within horizontal triplets, responses were slower in the impeded condition (M = 958 ms, SD = 229 ms) than in either the solid barrier (M = 911 ms, SD = 226 ms) and solid nonbarrier conditions (M = 896 ms, SD = 213 ms, ts > 2.42, ps < .05), for impeded versus solid barrier (d = .20, 95% CI [.03, .37]), for impeded versus solid nonbarrier (d = .27, 95% CI [.13, .42]). However, there were no differences within horizontal triplets between the impeded condition and either the pattern barrier (M = 1,020 ms, SD = 294 ms) or pattern nonbarrier conditions (M = 976 ms, SD = 305 ms, ts < 1.72, ps ≥ .10). Within vertical triplets, responses were slower in the impeded condition (M = 1,109 ms, SD = 260 ms) than in either the solid nonbarrier (M = 903 ms, SD = 246 ms) or solid barrier conditions (M = 895 ms, SD = 226 ms, ts > 13.63, ps < .001), for impeded versus solid barrier (d = .83, 95% CI [.68, .97]), for impeded versus solid nonbarrier (d = .79, 95% CI [.66, .93]). Within vertical triplets, responses were also slower for the impeded condition than in either the pattern barrier (M = 987 ms, SD = 292 ms) or pattern nonbarrier conditions (M = 1,005 ms, SD = 361 ms, ts > 2.46, ps < .05), for impeded versus pattern barrier (d = .43, 95% CI [.19, .67]), for impeded versus pattern nonbarrier (d = .29, 95% CI [.05, .53]).

Our second analysis asked whether the type and position of the distracting element influenced response times. We conducted a 2 (type: pattern vs. solid) × 2 (position: barrier vs. nonbarrier) within-subjects ANOVA, omitting the impeded condition. This analysis revealed that responses were faster for solid (M = 902 ms, SD = 238 ms) than for pattern trials (M = 997 ms, SD = 327 ms) as evidenced by a main effect of distractor type, $F(1, 21) = 11.92$, $p = .002$, $f = .47$. Response times did not differ between nonbarrier (M = 945 ms, SD = 301 ms) and barrier trials (M = 953 ms, SD = 279 ms), $F(1, 21) = 0.36$, $p = .56$, $f = .04$. There was also no evidence of an interaction between position and distractor type, $F(1, 21) = 0.12$, $p = .73$, $f = .02$.

Accuracy. To explore accuracy, proportion errors within each stimulus condition were averaged within each subject and we then conducted a 2 (Triplet) × 2 (Concordance) × 5 (Condition) within-subjects ANOVA on the proportion of errors. This analysis revealed a significant effect of condition, $F(4, 84) = 6.22$, $p < .001$, $f = 1.35$.

Participants made more errors in the impeded condition (M = .09, SD = .12) relative to the two solid conditions: solid barrier (M = .04, SD = .07), solid nonbarrier (M = .04, SD = .08, ts > 3.18, ps < .005); for impeded versus solid barrier (d = .50, 95% CI [.20, .80]), for impeded versus solid nonbarrier (d = .51, 95% CI [.17, .85]) but not between the impeded and the two pattern conditions: pattern barrier ($M = .10$, $SD = .16$), pattern nonbarrier ($M = .08$, $SD = .15$, $ts < .67$, $ps > .51$). There were no differences between the pattern nonbarrier and pattern barrier conditions ($t = 1.73$, $p = .10$) or the solid barrier and solid nonbarrier conditions ($t = .22$, *ns*).

To discover whether the type and position of the distracting element influenced error rates, we next conducted a 2 (type: pattern vs. solid) × 2 (position: barrier vs. nonbarrier) within-subjects ANOVA, omitting the impeded condition. This analysis revealed that participants made fewer errors on solid ($M = .04$, $SD = .07$) versus pattern trials ($M = .09$, $SD = .16$) as evidenced by a main effect of distractor type, $F(1, 21) = 9.15$, $p = .006$, $f = .31$. Error rates were not different between nonbarrier ($M = .06$, $SD = .12$) and barrier trials ($M = .07$, $SD = .13$), $F(1, 21) = 1.17$, $p = .29$, $f = .03$. There was also no evidence of an interaction between position and distractor type, $F(1, 21) = 2.14$, $p = .16$, $f = .04$.

Summary. The results of Experiment 2 support the idea that the low performance in the impeded condition seen in Experiment 1 stemmed specifically from the presence of competing correspondences. First, accuracy was lower, and response times slower, for pairs in impeded placement than for pairs that had a solid visual barrier between them. Thus, the adverse effects of impeded placements cannot be attributed to simple visual blocking. Second, competing patterns—triplets like those being compared—were more adverse than solid blocks—further evidence that competing correspondences interfere with alignment. It did not matter whether these competing elements were placed between the triplets, or off to the side. These findings are consistent with the idea that competition among potentially corresponding elements is detrimental to efficient spatial alignment and that direct alignment facilitates visual comparison.

Experiment 3

Thus far, effects have been demonstrated with pairs of visualizations that had object matches as well as relational matches. In Experiment 3 we test the prediction that the spatial alignment principle applies in purely relational comparisons. To do this, we paired shape triplets with color triplets, thus removing object matches.

Method

Participants. Participants were 25 adults ($M = 20.51$ years, range = 18 – 28 years, 15 women, 10 men) from Northwestern University. One participant was excluded for having previously completed a related study.

Based on the resulting sample size and the experimental design, we conducted a sensitivity analysis using the same specifications and assumptions as in Experiment 1. This analysis indicated that the Experiment was 80% powered to detect an effect of Cohen's $f = .25$.

Conditions. The experiment followed a four placement (horizontal, Oblique 1, Oblique 2, vertical) × 2 Triplets (horizontal vs. vertical) × 2 Concordance (same vs. different), within-subjects design.

Materials and procedure. The materials and procedure for Experiment 3 were identical to those of Experiment 1, with the following exceptions. Every trial in Experiment 3 consisted of a shape triplet paired with a color triplet, ensuring that there were no object matches on any trials; participants could make comparisons solely based on relational patterns. Also, to encourage participants to view the stimuli as structures with parts, instead of as holistic single objects, the shapes or colors within each triplet were spaced 40 pixels apart (in Experiment 1, they were spaced 10 pixels apart).

The shapes and colors were identical to those used in Experiment 1. To create trials, three combinations of shape pairings and three combinations of color pairings were randomly selected from the 12 possible pairings. These were then used to create triplets of either shapes or colors (e.g., triangle–square–triangle or blue–red–blue). The order of color and shape triplets (which ones appeared at the topmost or leftmost position on the screen relative to the bottommost or rightmost position on the screen) was counterbalanced across trials. As in Experiment 1, there were two thirds different trials and one third same trials. The design was 2 Triplet (horizontal or vertical) × 4 Placement (vertical, horizontal, Oblique 1, and Oblique 2) × 2 Concordance, all within-subjects.

The procedure was as in Experiment 1 with two exceptions. Participants completed 12 practice trials (as opposed to six) before completing experimental trials, to ensure that participants understood the more difficult relational task. Second, as separate shape and color trials were eliminated, there was no need for blocking; participants completed all trials in random order.

Results

Response time. We averaged response times for correct trials within each condition and subject and then conducted a 2 (triplet) × 4 (placement) × 2 (concordance) within-subjects ANOVA. This analysis indicated that responses were faster for horizontal ($M = 1{,}144$ ms, $SD = 236$ ms) relative to vertical triplets ($M = 1{,}183$ ms, $SD = 231$ ms) as revealed by a significant main effect of triplet, $F(1, 23) = 14.64$, $p = .001$, $f = .56$. The predicted Triplet × Placement interaction was also significant, $F(3, 69) = 18.39$, $p < .001$, $f = .85$ (see Figure 4).

To explore the Triplet × Placement interaction, we grouped the conditions based on their spatial placement conditions. Within horizontal triplets, response times were faster for direct ($M = 1118$ ms, $SD = 218$ ms) relative to impeded trials ($M = 1{,}178$ ms, $SD = 238$ ms, $t(23) = 3.49$, $p = .002$, $d = .25$, 95% CI [.11, .40]) and marginally faster for direct relative to indirect trials, $t(23) = 1.91$, $p = .07$, $d = .10$, 95% CI [−.01, .20]. Response times were also marginally faster for indirect ($M = 1{,}139$ ms, $SD = 225$ ms) relative to impeded trials, $t(23) = 1.94$, $p = .07$, $d = .16$, 95% CI [−.01, .34]. Within vertical triplets, response times were faster for direct ($M = 1{,}130$ ms, $SD = 221$ ms) relative to both indirect ($M = 1{,}181$ ms, $SD = 217$ ms) and impeded trials ($M = 1{,}238$ ms, $SD = 217$ ms, $ts > 4.38$, $ps < .001$), for direct versus indirect ($d = .23$, 95% CI [.13, .34]), for direct versus impeded ($d = .49$, 95% CI [.33, .65]). Response times were also faster for indirect relative to impeded trials, $t(23) = 4.52$, $p < .001$, $d = .26$, 95% CI [.14, .37].

Accuracy. After averaging the errors within each condition within each subject, a 2 (triplet) × 4 (placement) × 2 (concordance) within-subjects ANOVA on the proportion of errors re-

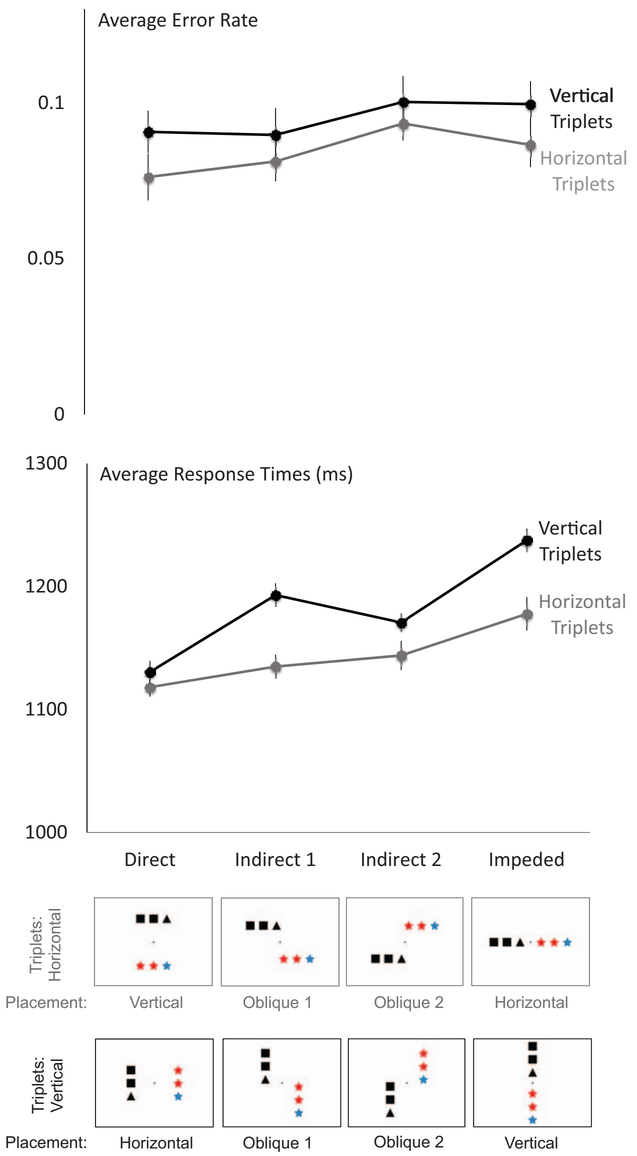

Figure 4. Results of Experiment 3 across Spatial Placement and Triplet conditions: average proportion of error (top panel) and average response time (in ms; bottom panel). Examples of each triplet and placement condition are shown at bottom. Error bars represent within-subject standard errors.

vealed that participants were more accurate for different (M = .07, SD = .09) relative to same trials (M = .11, SD = .12), as revealed by a significant main effect of concordance, $F(1, 23) = 8.91$, p = .007, f = .89. The Placement × Triplet interaction was not significant, $F(1, 69) = .56$, p = .64, f = .14. No other main effects or interactions were found (Fs < 2.10, ps > .15).

Summary. As predicted, even when comparisons must be made solely based on relational patterns, response time findings are consistent with the spatial alignment hypothesis. Specifically, we found that participants were faster for direct than for impeded trials, and—in vertical triplets—faster for indirect trials than for impeded trials.

Experiment 4

Experiments 1 and 2 showed the spatial alignment effect in visuals that contain object matches and Experiment 3 showed the spatial alignment effect in visuals that contain only relational matches. Comparing the effect sizes across these experiments suggests that spatial alignment effects are larger when visual comparisons contain object matches (i.e., spatial alignment effects were larger in Experiment 1 relative to Experiment 3). This finding, if robust, could have theoretical implications regarding the mechanism underlying the spatial alignment effect (see the General Discussion). Thus, the goal of Experiment 4 was to directly compare the effect sizes across visual comparisons with and without object matches in a within-subject manipulation. Moreover, Experiment 4 serves as an opportunity to replicate the effects observed in prior experiments.[1]

Method

Participants. An a priori power analysis for a F test on a repeated measures ANOVA with within-between interaction was conducted using G*Power 3.1, with the same assumptions as specified in Experiments 1 and 3. We aimed to detect between small and medium effects (Cohen's f = .15–.20), which revealed a required sample size of 36–62 participants. We recruited 65 adult participants from Northwestern University (M = 18.62 years, range = 18 – 22 years, 44 women, 21 men), anticipating some dropout. Ten participants were excluded due to incomplete data, leaving 55 participants in the analytic sample, which fell within the desired range.

Conditions. The conditions were the same as in Experiments 1 and 3, but with an additional factor to account for the comparison type—either visuals containing object matches as well as relational matches (the shapes version of Experiment 1) or visuals containing only relations (Experiment 3). Thus, the experiment followed a 4 Placement (horizontal, Oblique 1, Oblique 2, vertical) × 2 Triplets (horizontal versus vertical) × 2 Concordance (same vs. different) × 2 Comparison Type (Objects + Relations vs. Relations-only), within-subjects design. In addition, comparison type was presented in a blocked, counterbalanced order (see details to follow); thus, the presentation order (Objects + Relations first or Relations-only first) served as a between-subjects factor.

Materials and procedure. The materials were identical to Experiment 3 and the shape version of Experiment 1. As in prior experiments, participants were told to respond as fast and as accurately as possible and were encouraged to take a break between trials if they found themselves inattentive or drowsy.

Results

Trials with response times faster than 100ms were eliminated (<1% of trials). We next analyzed data separately by response time and error rates.

Response time. We averaged response times for correct trials within each condition and subject and then conducted a 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (comparison type) ×

[1] We thank an anonymous reviewer for this suggestion.

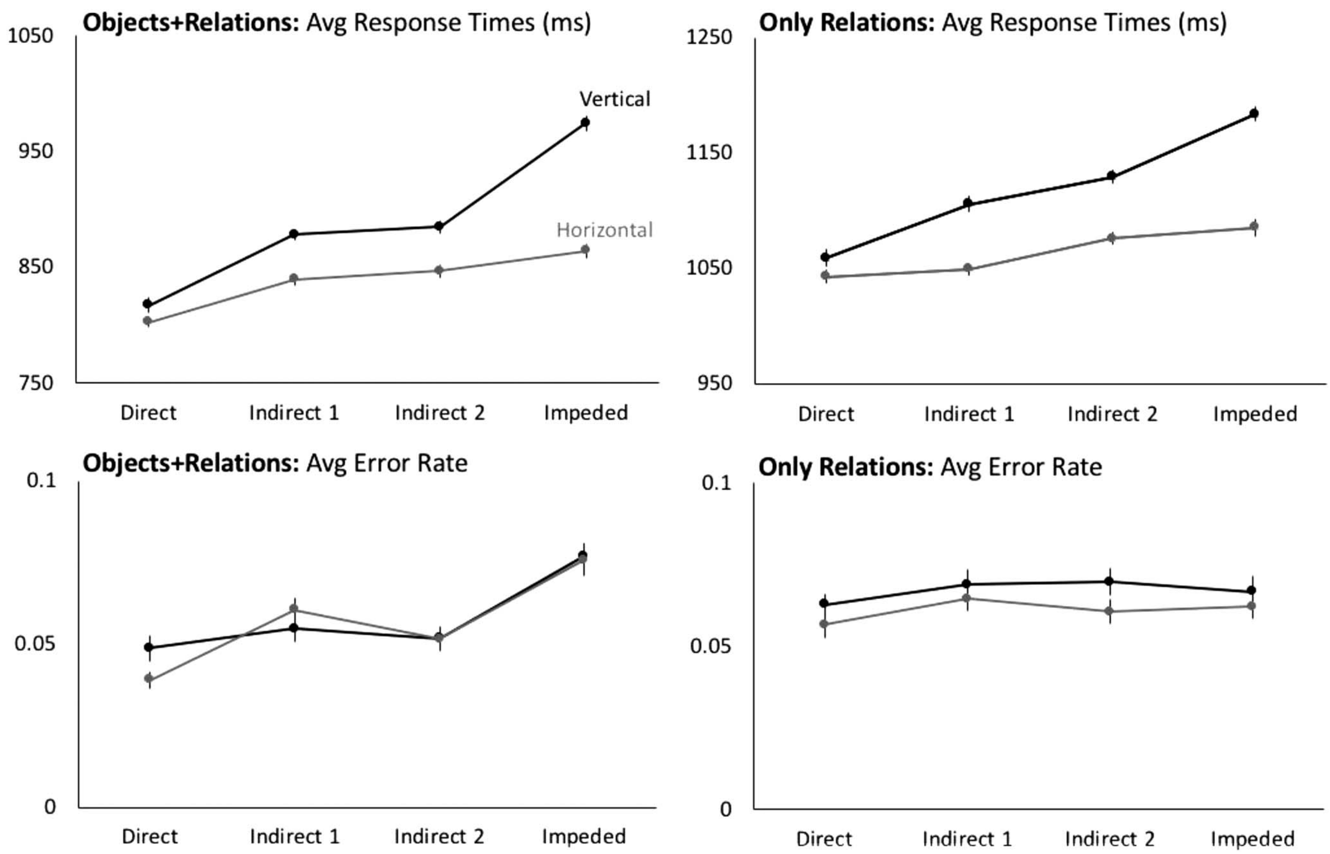

Figure 5. Results of Experiment 4 across spatial placement conditions for vertical triplets (black lines) and horizontal triplets (gray lines). Results are shown as (top left) average response time within the Objects + Relations condition; (top right) average response time within the Relations-only condition; (bottom left) average error rate within the Objects + Relations condition; (bottom right) average error rate within the Relations-only condition. Error bars represent within-subject standard errors.

2 (presentation order) mixed ANOVA. This analysis showed a significant Placement × Triplet × Comparison type interaction, $F(3, 159) = 3.63$, $p = .01$, $f = .27$. To explore this interaction, we reran the ANOVA separately within each comparison type.[2]

Objects + Relations condition. A 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (presentation order) mixed ANOVA showed that participants were faster for horizontal (M = 838 ms, SD = 143 ms) relative to vertical triplets (M = 888 ms, SD = 168 ms), as evidenced by a main effect of triplet, $F(1, 53) = 64.38$, $p < .001$, $f = .97$, and faster for same (M = 838 ms, SD = 141 ms) relative to different trials (M = 888 ms, SD = 169m s), as evidenced by a main effect of concordance, $F(1, 53) = 11.80$, $p = .001$, $f = .95$. There was also a significant main effect of placement $F(3, 159) = 17.38$, $p < .001$, $f = .65$. This analysis also revealed significant interactions between Triplet × Presentation order, $F(1, 53) = 5.12$, $p = .03$, $f = .27$, Concordance × Presentation order, $F(1, 53) = 4.29$, $p = .04$, $f = .57$, Placement × Concordance, $F(3, 159) = 7.47$, $p < .001$, $f = .37$, and the predicted Placement × Triplet interaction, $F(3, 159) = 100.55$, $p < .001$, $f = 1.50$.

To explore the Triplet × Placement interaction, we grouped the conditions based on their spatial placement conditions (see Figure 5). Within horizontal triplets, response times were faster for direct (M = 802 ms, SD = 126 ms) relative to both indirect (M = 843 ms, SD = 115 ms) and impeded trials (M = 864 ms, SD = 134 ms, ts > 6.57, ps < .001), for direct versus indirect (d = .33, 95% CI [.23, .43]), for direct versus impeded (d = .47, 95% CI [.33, .61]). Response times were also faster for indirect relative to impeded trials, $t(54) = 2.59$, $p = .01$, $d = .16$, 95% CI [.04, .29]. Within vertical triplets, response times were faster for direct (M = 817 ms, SD = 132 ms), relative to both indirect (M = 881 ms, SD = 130 ms) and impeded trials (M = 974 ms, SD = 150 ms, ts > 7.02, ps < .001), for direct versus indirect (d = .49, 95% CI [.34, .63]), for direct versus impeded (d = 1.10, 95% CI [.88, 1.31]). Response times were also faster for indirect relative to impeded trials, $t(54) = 10.54$, $p < .001$, $d = .64$, 95% CI [.51, .77].

Relations-only condition. A 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (presentation order) mixed ANOVA showed that participants were faster when the Relations-only condition followed the Objects + Relations condition (M = 1,013 ms, SD = 188 ms), relative to when it preceded the Object + Relations

[2] Additional main effects and interactions are described in the online supplementary materials.

condition (M = 1,161 ms, SD = 209 ms) as evidenced by a significant main effect of order $F(1, 53) = 9.23$, $p = .004$, $f = 2.40$. This finding is consistent with studies in which processing concrete pairs that share objects and relations facilitates the later processing of purely relational pairs (Kotovsky & Gentner, 1996; Thompson & Opfer, 2010). The analysis also showed that participants were faster for horizontal (M = 1,063 ms, SD = 206 ms) relative to vertical triplets (M = 1,119 ms, SD = 217 ms), as evidenced by a significant main effect of triplet, $F(1, 53) = 131.77$, $p < .001$, $f = .91$, and a significant main effect of placement, $F(3, 159) = 15.01$, $p < .001$, $f = .56$. This analysis also revealed a significant interaction between Placement × Triplet, $F(3, 159) = 44.34$, $p < .001$, $f = .96$.

To explore the predicted Triplet × Placement interaction, we grouped the conditions based on their spatial placement conditions (see Figure 5). Within horizontal triplets, response times were faster for direct (M = 1,042 ms, SD = 189 ms) relative to both indirect (M = 1,063 ms, SD = 189 ms) and impeded trials (M = 1,085 ms, SD = 219 ms, $ts > 2.74$, $ps < .01$), for direct versus indirect (d = .11, 95% CI [.03, .18]), for direct versus impeded (d = .20, 95% CI [.08, .31]). Response times were also faster for indirect relative to impeded trials, $t(54) = 2.21$, $p = .03$, (d = .10, 95% CI [.01, .19]). Within vertical triplets, response times were faster for direct (M = 1,059 ms, SD = 208 ms) relative to both indirect (M = 1,117 ms, SD = 197 ms) and impeded trials (M = 1,184 ms, SD = 199 ms, $ts > 5.91$, $ps < .001$), for direct versus indirect (d = .29, 95% CI [.19, .38]), for direct versus impeded (d = .61, 95% CI [.49, .73]). Response times were also faster for indirect relative to impeded trials, $t(54) = 8.13$, $p < .001$, $d = .34$, 95% CI [.25, .42].

Effects across comparison types. We next examined effects of spatial placement across comparison types (Objects + Relations vs. Relations-only). Figure 6 summarizes the effect size differences in response times across spatial placement, comparison type, and triplet conditions for Experiments 1, 3, and 4.

Several patterns are discernable from Figure 6. First, observed effects are consistently larger for the objects + relations condition relative to the Relations-only condition—this is especially the case in vertical triplets. Second, effects in the present experiment (Experiment 4) replicate effects observed in earlier experiments (Experiments 1 and 3). Third, effect sizes tend to be larger for vertical triplets overall, relative to horizontal triplets. Finally, effects tend to be largest for direct versus impeded comparisons.

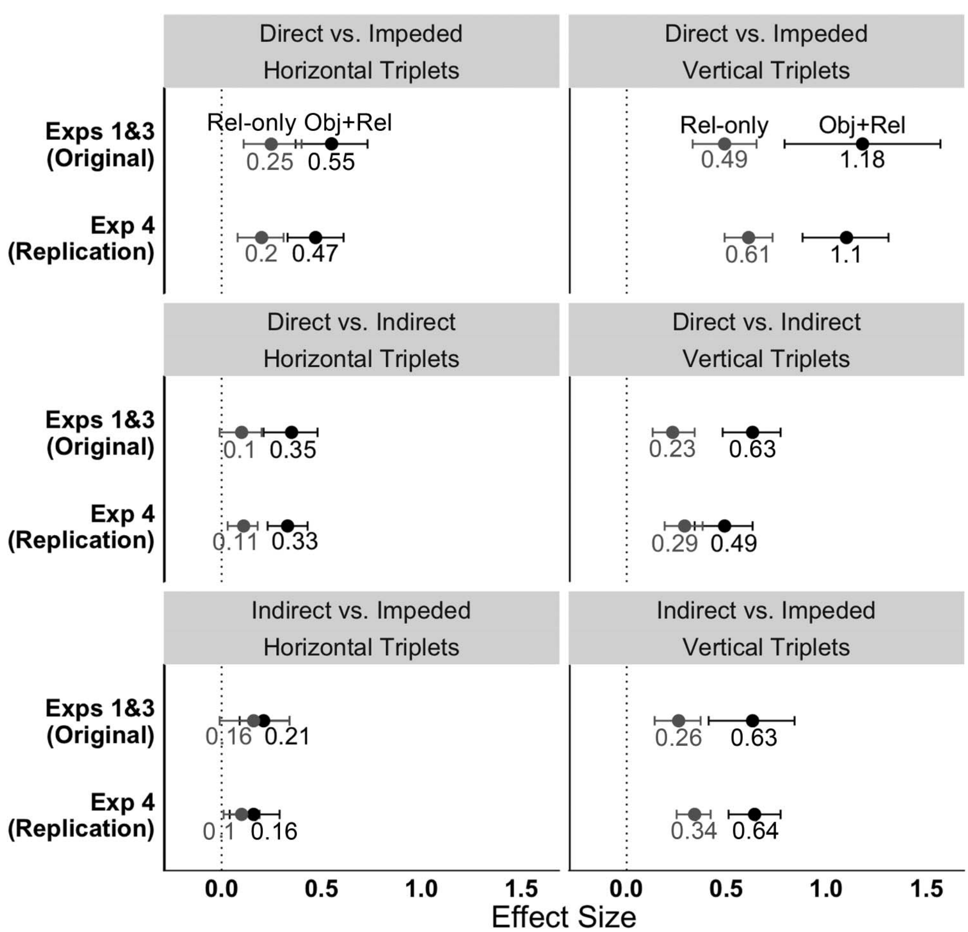

Figure 6. Effect sizes (in Cohen's d) across Experiments 1, 3, and 4 for each spatial placement comparison (direct vs. impeded, direct vs. indirect, and indirect vs. impeded), triplet condition (vertical or horizontal), and comparison type (Objects + Relations or Relations-only).

To directly examine the relationship between these conditions, we conducted a 2 (comparison type) × 3 (spatial placement) × 2 (triplet) within-subjects ANOVA. This analysis revealed significant main effects of triplet, $F(1, 54) = 161.95$, $p < .001$, $f = .37$, comparison type $F(1, 54) = 146.75$, $p < .001$, $f = 1.98$, and spatial placement, $F(2, 108) = 130.82$, $p < .001$, $f = .47$, qualified by significant interactions between Triplet × Spatial Placement, $F(2, 108) = 40.45$, $p < .001$, $f = .22$, and Comparison Type × Spatial Placement, $F(2, 108) = 4.30$, $p = .02$, $f = .06$.

Because the Comparison Type × Spatial Placement interaction was of theoretical interest, we further examined this interaction by conducting paired *t* tests on the average differences between the comparison types for each spatial placement comparison (direct vs. impeded, direct vs. indirect, and indirect vs. impeded alignments). This analysis revealed that (1) the difference in response times between the direct versus impeded placements was statistically larger for the Objects + Relations condition ($M = -110$ ms, $SD = 57$), relative to the Relations-only condition ($M = -84$, $SD = 63$, $t = 2.91$, $p = .005$, $d = .43$, 95% CI [.12, .74]); (2) the difference in response times between the direct versus indirect comparisons was marginally larger for the Objects + Relations ($M = -52$, $SD = 43$), relative to the Relations-only condition ($M = -39$, $SD = 48$, $t = 1.77$, $p = .08$, $d = .29$, 95% CI [−.04, .62]); and (3) the difference in response times between the indirect versus impeded placements was not statistically larger for the Objects + Relations ($M = -57$, $SD = 48$), relative to the Relations-only condition ($M = -45$, $SD = 46$, $t = 1.63$, $p = .11$, $d = .27$, 95% CI [−.07, .61]). Thus, this analysis suggests that the larger spatial placement differences observed in Objects + Relations versus Relations-only comparisons is primarily driven by the response time differences in direct versus impeded spatial placements.

Accuracy. To examine participant accuracy, we averaged error rates within each condition and subject and then conducted a 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (comparison type) × 2 (presentation order) mixed ANOVA. This analysis showed a significant Placement × Triplet × Comparison Type interaction, $F(3, 159) = 5.90$, $p = .001$, $f = .37$. To explore this interaction, we reran the ANOVA separately within each comparison type.

Objects + Relations condition. A 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (presentation order) mixed ANOVA showed significant interactions between Placement × Order, $F(3, 159) = 3.87$, $p = .01$, $f = .26$, the predicted Placement × Triplet interaction, $F(3, 159) = 17.32$, $p < .001$, $f = .60$. The Placement × Triplet interaction was qualified by a significant interaction between Placement × Triplet × Concordance, $F(3, 159) = 4.63$, $p = .004$, $f = .30$. To explore this further, we conducted 2 (triplet) × 4 (placement) × 2 (presentation order) mixed ANOVAs separately within same and different trials. This analysis revealed that, in both cases, the Placement × Triplet interaction was significant, but that this interaction was slightly stronger in same, $F(3, 159) = 11.67$, $p < .001$, $f = .47$, versus different trials, $F(3, 159) = 8.99$, $p < .001$, $f = .41$.

Because the Placement × Triplet interaction was significant in both concordance conditions, we proceeded to explore effects by grouping by spatial placement conditions (see Figure 5). Within horizontal triplets, error rates were lower for direct ($M = .04$, $SD = .07$) relative to both indirect ($M = .06$, $SD = .06$) and impeded trials ($M = .08$, $SD = .08$, ts > 4.49, ps $< .001$), for direct versus indirect ($d = .26$, 95% CI [.14, .38]), for direct versus impeded ($d = .53$, 95% CI [.34, .72]). Error rates were also lower for indirect relative to impeded trials, $t(54) = 2.77$, $p = .008$; ($d = .30$, 95% CI [.08, .51]). Within vertical triplets, error rates were lower for direct ($M = .05$, $SD = .07$) relative to impeded trials ($M = .08$, $SD = .06$), $t(54) = 4.62$, ($p < .001$ $d = .44$, 95% CI [.24, .63]), but not for direct relative to indirect trials (*ns*, $d = .07$, 95% CI [−.12, .26]). Indirect trials ($M = .05$, $SD = .06$) exhibited lower error rates relative to impeded trials, $t(54) = 3.93$, $p < .001$; $d = .41$, 95% CI [.19, .62].

Relations-only condition. A 2 (triplet) × 4 (placement) × 2 (concordance) × 2 (presentation order) mixed ANOVA showed only a significant main effect of presentation order $F(3, 53) = 9.76$, $p = .003$, $f = 1.24$; participants had lower error rates when the Relations-only condition preceded the object-match condition ($M = .04$, $SD = .07$), relative to when it followed the object-match condition ($M = .09$, $SD = .11$). The Placement × Triplet interaction was not statistically significant, $F(3, 159) = 0.47$, $p = .70$, $f = .11$.

To explore planned comparisons, we grouped the conditions based on their spatial placement conditions (see Figure 5). Within horizontal triplets, error rates were not statistically different between direct ($M = .06$, $SD = .07$), indirect ($M = .06$, $SD = .06$), or impeded trials ($M = .06$, $SD = .06$, ts <1.04, ps $> .30$, ds $\leq .10$). Within vertical triplets, error rates were not statistically different between direct ($M = .06$, $SD = .07$), indirect ($M = .07$, $SD = .06$), or impeded trials ($M = .07$, $SD = .07$, ts < 1.46, ps $> .15$, ds ≤ 10).

Summary. The response time results of Experiment 4 followed spatial alignment predictions in both visual comparisons that contained object-matches (replicating Experiment 1) and in visual comparisons that did not contain object matches (only relations; replicating Experiment 3). Moreover, spatial alignment was found to facilitate participant accuracy in visual comparisons with object-matches, but not in visual comparisons with only relations (findings that were also consistent Experiments 1 and 3). A further contribution of Experiment 4 was that it provided a direct comparison between the size of the spatial alignment effects between visual comparisons with and without object matches. Here we found that spatial alignment effects were larger in comparisons containing object matches. We speculate about the implications of this finding in the general discussion.

General Discussion

Our experiments suggest spatial alignment as a new principle of effective visual comparison. There are three key results that support this idea: (1) in all four studies, direct spatial placement resulted in more efficient visual comparison than impeded spatial placement; (2) the impedance effect was specific to structurally similar items, which offered potential competing correspondences (Experiment 2)—it did not occur for simple physical barriers; and (3) the advantage of direct over impeded placements was found for purely relational pairs, for which direct object-matching cannot apply (Experiments 3 and 4).

Direct spatial correspondence between matching objects almost certainly facilitates the process needed for judging identity matches, as in Experiment 1. But critically, this account cannot explain the results of Experiment 3, in which there were no actual object matches. Instead, the similarity was one of relational patterns. Participants in this study had to respond "same" to pairs such as square–square–circle and red–red–blue. In order to see that an individual shape corresponds to an individual color, participants had to match the patterns, not the elements. Yet we still found an advantage of direct spatial alignment in facilitating purely relational comparisons.

Many researchers have proposed that analogical comparison processes are used to process visual comparisons for visuals that, like the ones used here, contain elements in spatial configurations (Doumas et al., 2008; Gattis, 2002; Kotovsky & Gentner, 1996; Lovett & Forbus, 2017; Sagi et al., 2012). Under the structure-mapping account of analogical comparison, components are placed into correspondence based on aligning common relational structure (e.g., Gentner, 1983, 2010). On this account, the direct alignment advantage in speed and accuracy arises because it maximizes the clarity with which the relevant matches can be found and minimizes the presence of close competing potential matches. It also explains that the impedance effect is due to the presence of competing correspondences, and not simply to visual barriers as found in Experiment 2. Finally, it explains that the advantage of direct over impeded alignment should hold for purely relational matches, which lack concrete matches of objects and properties, as found in Experiments 3 and 4.

The results may also be explained by a simpler account of how the visual system might represent the positions of the objects as spatial patterns. When encoding a pattern of letters or colors with a horizontal structural axis, the horizontal dimension of visual space holds the most critical information about sequence of object identities across the pattern (Ragni & Knauff, 2013; see also Franconeri, Scimeca, Roth, Helseth, & Kahn, 2012 for a potential mechanism for coding the internal relations across this sequence). Arranging two such sets vertically may allow the visual system to use the horizontal and vertical dimensions to independently code the pattern (horizontal) and the axis that separates the triplets (vertical). This avoids using the same representation of a dimension for both the structural and arrangement axes, which could cause strong interference between or overwriting of either representation (Franconeri, Alvarez, & Cavanagh, 2013), leading to errors and/or slower processing. Oblique arrangement could present a case of partial interference. In Experiment 3—where no object matches are available—viewers could compare abstracted perceptual groupings. For example, a horizontal AAB pattern might be abstractly represented as a spatial pattern of two items that perceptually group, followed by one on the right that does not (Yu, Tam, & Franconeri, 2019). That spatial pattern would be the same for both an AAB pattern in shape or color, allowing a same-different relational comparison of the two (e.g., Huang & Pashler, 2007).

Limitations and Unexplained Results

The predictions of the spatial alignment hypothesis were supported in response times for both Objects + Relations and Relations-only comparisons. However, for error rates, spatial alignment predictions were only supported in the Objects + Relations condition—for purely relational pairs, error rates were not different between the alignment conditions. Further, although response time predictions were borne out in purely relational pairs, the effect sizes were much larger for pairs that include object matches as well as relational matches. Though this set of experiments did not attempt to tease apart an analogical versus a perceptual account for explaining performance on these two comparison types, these findings are generally consistent with analogical models of the online processing of comparisons, which assume that both object matches and relational matches enter into the mapping process (Doumas et al., 2008; Falkenhainer, Forbus, & Gentner, 1989; Holyoak & Thagard, 1989; Kokinov & French, 2003; Thibaut et al., 2010; see also Goldstone & Medin, 1994). In the analogical framework, this is explained as follows: (1) in Relations-only pairs, same–different responses must be made on the basis of matching (or mismatching) relational patterns, whereas in Objects + Relations matches, same–different responses can also be made on the basis of matching (or mismatching) objects, and (2) object matches and mismatches are much easier to detect in direct than in impeded placement. In general, object matches (and mismatches) are highly salient (Gentner & Toupin, 1986) and are detected more quickly than relational matches/mismatches (Gentner & Kurtz, 2006; Goldstone, Medin, & Gentner, 1991; Love, Rouder, & Wisniewski, 1999). This is reflected in the fact that response times overall are far lower for Objects + Relations pairs than for Relations-only pairs (see Figure 5).

An unexpected finding concerns the faster performance for horizontal triplets in horizontal (impeded) placements relative to vertical triplets in vertical (impeded) placements. Although this finding is not inconsistent with the spatial alignment hypothesis, it is not predicted by it. We suspect that the advantage for horizontal placement stems from extensive practice in reading, resulting in highly fluent encoding of horizontal relational sequences (see Thibaut, French, & Vezneva, 2010, for a related explanation). Though this hypothesis is purely speculative, it remains a possibility that could be examined in future research.

Implications for Education and Design

Visualizations play a critical role in education, in which spatial patterns are common (Ainsworth et al., 2011; Forbus, Usher, Lovett, Lockwood, & Wetzel, 2011; Gattis & Holyoak, 1996; Jee et al., 2014; Kellman, 2013; Uttal et al., 2006). For example, in medical contexts, visual comparisons have been shown to be helpful in learning to distinguish between diseased and heathy lungs (Kok et al., 2013) and in diagnosing skin diseases (Brooks et al., 1991; Kellman, 2013). Visual comparisons are also prominent in depictions of economic and political patterns that need to be considered in policy decisions, as well as in education (e.g., Gentner et al., 2016; Kok et al., 2013; Matlen et al., 2011; Yuan et al., 2017). Yet not all such visualizations are equally effective.

The spatial alignment principle may explain why some visual comparisons are more effective than others. For example, spatial alignment can explain why matching legend entries to lines in a graph is more efficient when the legend is placed in direct spatial alignment to the lines (Wong, 2013; see the first column in Figure 1). Direct spatial alignment also facilitates quantitative decisions, such as determining which of two bars in a bar graph is longer (Cleveland & McGill, 1985, 1987; see Figure 7). More generally, as shown in Experiment 2, the presence of

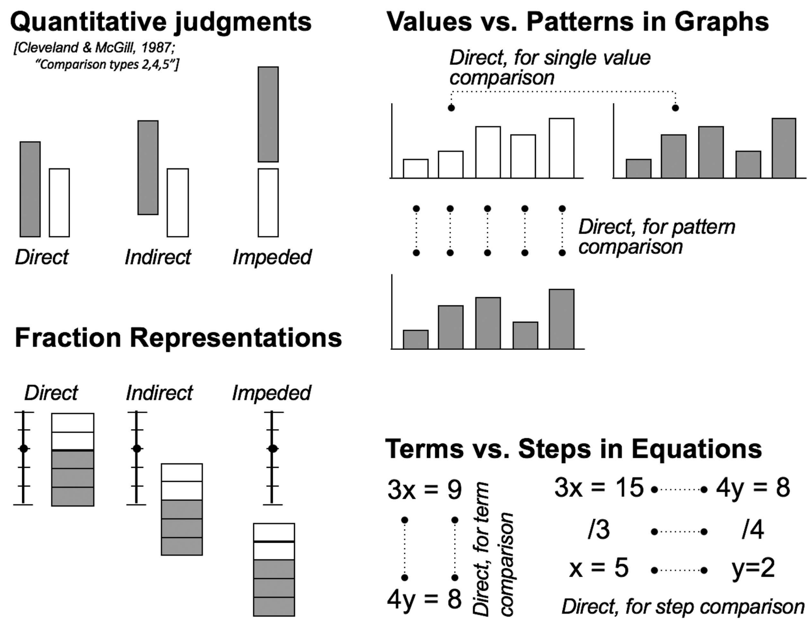

Figure 7. Examples of direct and impeded placements across different domains.

competing potential correspondences can result in less efficient comparison.

These findings also provide recommendations for designing instructional visualizations. Specifically, visual comparisons could be organized based on which relations are important to align. For example, when the goal is to teach students a series of solution steps, the optimal placement of a series of equations would be in side-by-side fashion, allowing students to align the order of solution steps (see Figure 7, bottom right). But if instead the goal is to show students the parallel between $3x = 9$ and $4y = 8$, then the optimal alignment would be in a vertical placement. Another obvious implication is that nonessential competing elements should be avoided to the extent possible. By making visual comparisons easier to process we can enhance our ability to convey visual information efficiently, supporting learning in science and mathematics.

We suggest that the spatial alignment principle provides a principled basis by which visualizations can be arranged for maximally efficient and informative comparisons. Given the importance of visual comparisons in education and decision-making, adhering to this principle could have an immediate impact on learning from visualizations.

References

Ainsworth, S., Prain, V., & Tytler, R. (2011). Science education. Drawing to learn in science. *Science, 333,* 1096–1097. http://dx.doi.org/10.1126/science.1204153

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48,* 87–113. http://dx.doi.org/10.1080/00461520.2013.775712

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General, 120,* 278–287. http://dx.doi.org/10.1037/0096-3445.120.3.278

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11,* 356–373. http://dx.doi.org/10.1080/15248371003700015

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science, 229,* 828–833. http://dx.doi.org/10.1126/science.229.4716.828

Cleveland, W. S., & McGill, R. (1987). Graphical perception: The visual decoding of quantitative information on graphical displays of data. *Journal of the Royal Statistical Society Series A (General), 150,* 192–229. http://dx.doi.org/10.2307/2981473

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1,* 42–45. http://dx.doi.org/10.20982/tqmp.01.1.p042

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

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

Farell, B. (1985). Same-different judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin, 98,* 419–456. http://dx.doi.org/10.1037/0033-2909.98.3.419

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2011). CogSketch: Sketch understanding for cognitive science research and for education. *Topics in Cognitive Science, 3,* 648–666. http://dx.doi.org/10.1111/j.1756-8765.2011.01149.x

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences, 17,* 134–141. http://dx.doi.org/10.1016/j.tics.2013.01.010

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. http://dx.doi.org/10.1016/j.cognition.2011.11.002

Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change.

Learning and Instruction, 22, 47–61. http://dx.doi.org/10.1016/j.learninstruc.2011.06.002

Gattis, M. (2002). Structure mapping in spatial reasoning. *Cognitive Science, 17,* 1157–1183. http://dx.doi.org/10.1016/S0885-2014(02)00095-3

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 231–239.

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

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

Gentner, D., & Kurtz, K. J. (2006). Relations, objects, and the composition of analogies. *Cognitive Science, 30,* 609–642. http://dx.doi.org/10.1207/s15516709cog0000_60

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52,* 45–56. http://dx.doi.org/10.1037/0003-066x.52.1.45

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2016). Rapid learning in a children's museum via analogical comparison. *Cognitive Science, 40,* 224–240. http://dx.doi.org/10.1111/cogs.12248

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14,* 487–513. http://dx.doi.org/10.1016/S0885-2014(99)00016-7

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10,* 277–300. http://dx.doi.org/10.1207/s15516709cog1003_2

Goldstone, R. L., & Medin, D. L. (1994). Time course of comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 29–50. http://dx.doi.org/10.1037/0278-7393.20.1.29

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology, 23,* 222–262. http://dx.doi.org/10.1016/0010-0285(91)90010-L

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32,* 717–742. http://dx.doi.org/10.1006/jmla.1993.1036

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295–355. http://dx.doi.org/10.1207/s15516709cog1303_1

Hribar, A., Haun, D. B. M., & Call, J. (2012). Children's reasoning about spatial relational similarity: The effect of alignment and relational complexity. *Journal of Experimental Child Psychology, 111,* 490–500. http://dx.doi.org/10.1016/j.jecp.2011.11.004

Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review, 114,* 599–631. http://dx.doi.org/10.1037/0033-295X.114.3.599

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1140–1160. http://dx.doi.org/10.1037/a0015019

Jee, B., Gentner, D., Uttal, D., Sageman, B., Forbus, K., Manduca, C., . . . Tikoff, B. (2014). Drawing on experience: How domain knowledge is reflected in sketches of scientific structures and processes. *Research in Science Education, 44,* 859–883. http://dx.doi.org/10.1007/s11165-014-9405-2

Jones, M., & Love, B. C. (2007). Beyond common features: The role of roles in determining similarity. *Cognitive Psychology, 55,* 196–231. http://dx.doi.org/10.1016/j.cogpsych.2006.09.004

Kellman, P. J. (2013). Adaptive and perceptual learning technologies in medical education and training. *Military Medicine, 178*(Suppl. 10), 98–106. http://dx.doi.org/10.7205/MILMED-D-13-00218

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science, 2,* 285–305. http://dx.doi.org/10.1111/j.1756-8765.2009.01053.x

Kok, E. M., de Bruin, A. B., Robben, S. G., & van Merriënboer, J. J. (2013). Learning radiological appearances of diseases: Does comparison help? *Learning and Instruction, 23,* 90–97. http://dx.doi.org/10.1016/j.learninstruc.2012.07.004

Kokinov, B., & French, R. (2003). Computational models of analogy-making. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 113–118). London, UK: Nature Publishing Group.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67,* 2797–2822. http://dx.doi.org/10.1111/j.1467-8624.1996.tb01889.x

Krawczyk, D. C., Holyoak, K. J., & Hummel, J. E. (2005). The one-to-one constraint in analogical mapping and inference. *Cognitive Science, 29,* 797–806. http://dx.doi.org/10.1207/s15516709cog0000_27

Kurtz, K. J., & Gentner, D. (2013). Detecting anomalous features in complex stimuli: The role of structured comparison. *Journal of Experimental Psychology: Applied, 19,* 219–232. http://dx.doi.org/10.1037/a0034395

Kurtz, K. J., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10,* 417–446. http://dx.doi.org/10.1207/S15327809JLS1004new_2

Larsen, A., & Bundesen, C. (1998). Effects of spatial separation in visual pattern matching: Evidence on the role of mental translation. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 719–731. http://dx.doi.org/10.1037/0096-1523.24.3.719

Larsen, A., McIlhagga, W., & Bundesen, C. (1999). Visual pattern matching: Effects of size ratio, complexity, and similarity in simultaneous and successive matching. *Psychological Research, 62,* 280–288. http://dx.doi.org/10.1007/s004260050058

Love, B. C., Rouder, J. N., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology, 38,* 291–316. http://dx.doi.org/10.1006/cogp.1998.0697

Lovett, A., & Forbus, K. (2017). Modeling visual problem solving as analogical reasoning. *Psychological Review, 124,* 60–90. http://dx.doi.org/10.1037/rev0000039

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467. http://dx.doi.org/10.1006/cogp.1993.1011

Matlen, B. (2019). *Spatial Alignment*. Retrieved from osf.io/sk9vt

Matlen, B. J., Vosniadou, S., Jee, B., & Ptouchkina, M. (2011). Enhancing the comprehension of science text through visual analogies. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2910–2915). Retrieved from https://escholarship.org/content/qt6fq355bv/qt6fq355bv.pdf

Paik, J. H., & Mix, K. S. (2008). It's all relative: Different levels of relational similarity used in children's comparisons. *British Journal of Developmental Psychology, 26,* 499–505. http://dx.doi.org/10.1348/026151007X260163

Ragni, M., & Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychological Review, 120,* 561–588. http://dx.doi.org/10.1037/a0032460

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review, 29,* 717–761. http://dx.doi.org/10.1007/s10648-016-9365-3

R Core Team. (2018). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99,* 561–574. http://dx.doi.org/10.1037/0022-0663.99.3.561

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36,* 1019–1050. http://dx.doi.org/10.1111/j.1551-6709.2012.01250.x

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. http://dx.doi.org/10.1016/j.jecp.2010.01.001

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81,* 1768–1786. http://dx.doi.org/10.1111/j.1467-8624.2010.01509.x

Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science, 3,* 499–535. http://dx.doi.org/10.1111/j.1756-8765.2010.01113.x

Uttal, D. H., Fisher, J. A., & Taylor, H. A. (2006). Words and maps: Developmental changes in mental models of spatial information acquired from descriptions and depictions. *Developmental Science, 9,* 221–235. http://dx.doi.org/10.1111/j.1467-7687.2006.00481.x

Wong, D. M. (2013). *The Wall Street Journal guide to information graphics: The dos and don'ts of presenting data, facts, and figures*. New York, NY: WW Norton.

Yu, D., Tam, D., & Franconeri, S. L. (2019). Gestalt similarity groupings are not constructed in parallel. *Cognition, 182,* 8–13. http://dx.doi.org/10.1016/j.cognition.2018.08.006

Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. *Developmental Psychology, 53,* 1098–1114. http://dx.doi.org/10.1037/dev0000302

Received July 16, 2018
Revision received December 18, 2019
Accepted December 19, 2019 ■