

RESEARCH ARTICLE

Are Categorical Spatial Relations Encoded by Shifting Visual Attention between Objects?

Lei Yuan^{1*}, David Uttal, Steven Franconeri

Department of Psychology, Northwestern University, Evanston, Illinois, United States of America

[‡] Current address: Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana, United States of America

* leiyuan@indiana.edu



OPEN ACCESS

Citation: Yuan L, Uttal D, Franconeri S (2016) Are Categorical Spatial Relations Encoded by Shifting Visual Attention between Objects? PLoS ONE 11 (10): e0163141. doi:10.1371/journal.pone.0163141

Editor: Ramesh Kumar Mishra, University of Hyderabad, INDIA

Received: December 22, 2015

Accepted: September 2, 2016

Published: October 3, 2016

Copyright: © 2016 Yuan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The name of our data repository is Dryad, and the Provisional doi is: [10.5061/dryad.7bc49](https://doi.org/10.5061/dryad.7bc49).

Funding: The research reported here was supported by the Institute of Education Sciences (<https://ies.ed.gov>), U.S. Department of Education, through Grant R305A120531 to Steven Franconeri, and by National Science Foundation (<http://www.nsf.gov>) Grant SBE-1041707 to David Uttal through the Spatial Intelligence and Learning Center (SILC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Perceiving not just values, but relations between values, is critical to human cognition. We tested the predictions of a proposed mechanism for processing categorical spatial relations between two objects—the *shift account* of relation processing—which states that relations such as ‘above’ or ‘below’ are extracted by shifting visual attention upward or downward in space. If so, then shifts of attention should improve the representation of spatial relations, compared to a control condition of identity memory. Participants viewed a pair of briefly flashed objects and were then tested on either the relative spatial relation or identity of one of those objects. Using eye tracking to reveal participants’ voluntary shifts of attention over time, we found that when initial fixation was on neither object, relational memory showed an absolute advantage for the object following an attention shift, while identity memory showed no advantage for either object. This result is consistent with the shift account of relation processing. When initial fixation began on one of the objects, identity memory strongly benefited this fixated object, while relational memory only showed a relative benefit for objects following an attention shift. This result is also consistent, although not as uniquely, with the shift account of relation processing. Taken together, we suggest that the attention shift account provides a mechanistic explanation for the overall results. This account can potentially serve as the common mechanism underlying both linguistic and perceptual representations of spatial relations.

Introduction

Humans are relational thinkers. Instead of relying on absolute values, much of cognition is based on comparing the relations among different entities [1–3]. For example, the perception of luminance is based on seeing contrast between two adjacent areas [4], decision-making is greatly influenced by establishing anchors and drawing comparisons [5], and even infants’ ability to judge length or height is based on comparison with other objects instead of absolute units [6]. Representing spatial relations is especially important to cognition, serving as the basis for our understanding of maps, graphs, and diagrams [7,8] and potentially underlying

Competing Interests: The authors have declared that no competing interests exist.

our understanding of seemingly abstract concepts through spatial metaphors for time, mood, or dominance [9–11].

Here we examine the ability to perceive novel categorical relations between two objects, e.g. seeing a book as being above a stapler on a shelf. We use the term ‘novel’ to exclude highly practiced spatial relationships that may be processed as experience-based templates, such as recognizing that a nose is above a mouth, a blue sky is over brown sand, or that a line graph reflects an interaction because it contains a prototypical ‘X’ shape [12–16]. We use the term ‘categorical’ because we focus on judgments of relative position (e.g., above/below) that abstract over distances between objects (e.g., two inches to the left)—such coordinate spatial relations may be processed by a different mechanism [17–19]. It is critical to understand how we process novel categorical relations between objects [20,21], because this class of relations is foundational to our ability to process maps, data visualizations, evolutionary trees, dental x-rays, mechanical diagrams, molecular structures, or algebraic equations [22–25].

Novel categorical relations are also interesting because they present a substantial challenge to our visual processing system. Evidence from visual search tasks—in which a viewer is asked to locate objects in a particular relation in a field of similar objects, such as a red bar to the left of a green bar—is consistent with a severe capacity limit, perhaps even of a single relation at a time [26–29]. Across these experiments, careful controls show that the root difficulty is the relational judgment itself, not the extraction of the identities of the objects: finding a red bar among green bars can be done in parallel, but determining the relations between two bars cannot. According to high-level accounts of relation perception, the root cause of this difficulty is a ‘binding problem,’ of tying each feature to its relative spatial location—there were red and green bars, but which one was above the other? [13,26,30].

How does the visual system solve this binding problem when processing novel categorical relations between objects? A partial solution is to filter processing for one location at a time—there is a red and a green, and attending to one of these locations reveals that it contains green [31–34]. But this filter does not reveal the *relative* location (above or below?) of that green with respect to the red [12,35,36]. We test the predictions of a specific proposal: that relative locations are extracted by shifting visual attention between the individual judged objects [12]. For the case of ‘above’ or ‘below,’ these shifts would be upward or downward in space. For our book and stapler example, shifting upward and landing on the book means that it is on the top. Shifting back down and landing on the stapler means that it is on the bottom. This attention shifting mechanism could solve the binding problem of linking object features to the objects that they belong to, by isolating processing of a single feature and a single object at a time. The shift of attention could also serve to compute the relative position of one of the objects in a categorical fashion, by encoding the direction of the shift, but ignoring the metric distance of the shift (see [12] for discussions of this mechanism, as well as alternative models). In support of this possibility, evidence from electrophysiological tracking of attention shows that attention does shift between objects during the simplest relational judgments [12].

This attention-shift mechanism has an idiosyncratic feature that makes a potentially falsifiable prediction: each shift only binds one ‘side’ of an asymmetric relational representation, and that two shifts would be needed to bind both sides. For vertical relations, neither the ‘above’ nor ‘below’ relation should be available without vertical shifts of attention. The upward shift lets you know that the book is on the top, and the downward shift that the stapler is on the bottom. This account leads to a counterintuitive prediction that logically equivalent spatial relationships, such as “A is above B” and “B is below A,” may be quite different at a perceptual level of representation.

This perceptual asymmetry would be congruent with a well-established asymmetry in the *linguistic* representation of relations [37–39]. While logically equivalent, *saying* that “A is

above B” versus “B is below A” means different things to a listener, by putting one object in a ‘target’ or ‘figure’ role and the other in a ‘referent’ or ‘ground’ role. For example, it sounds natural to say that a “bicycle is next to a building”, but would be odd to hear that “a building is next to a bicycle” [38]. Such asymmetries are not restricted to spatial relation descriptions and can be found throughout other forms of language [40,41].

Past work has attempted to link asymmetry in linguistic descriptions of relations to a similar potential asymmetry at the perceptual level, by using a sentence-picture verification task [42,43]. After seeing a sentence (e.g., “Star is above Line”), participants were asked if the sentence matched a picture of objects in that relation, or the opposite relation. People were faster at responding to questions that involved the word “above” than “below”, suggesting a preferred asymmetric way of representing the relation as “Star is above Line” instead of “Line is below Star”. This result, and many others, is consistent with the idea that perceptual representations of relations carry a similar type of asymmetric figure/ground relationship as language [44,45].

There is some evidence that this perceptual asymmetry might be driven by the way that people shift attention between objects during online processing. The preference for ‘above’ framings remained when participants were asked to pay attention to the top object [42]—which appears to be a default strategy [46]—but reversed when they were instructed to pay attention to the bottom object, or when the bottom object was made more relatively salient [43]. A more recent study using this type of sentence-picture verification task showed that when attention was more explicitly cued to one object in a relation, sentence verification was faster when the linguistic ‘figure’ object was cued [47]. Other studies have shown similar effects at the level of scene descriptions [48,49]. For instance, subliminally cueing one object in simple scenes can lead viewers to describe it as either “A dog chasing a man”, or “A man being chased by a dog” [50].

The current experiments

If the perception of novel categorical relations relies on an asymmetric representation, with “A above B” being different than “B below A”, then the attentional shift account of relation perception could explain the roots of this asymmetry. Shifting up produces “A is above”, and shifting down produces “B is below”. While previous evidence suggests such a role for attention, all of this evidence is based on congruency between perceived relations and linguistically described relations. While suggestive of an asymmetry in perception per se, it remains possible that the reported asymmetries in perceptual representations are due to participants linguistically encoded these relations (either implicitly or explicitly) [47].

Here we tested whether relational asymmetries are associated with patterns of attentional shifts, in a manner that isolates the perceptual level of representation, by employing a strong verbal interference concurrent task that should minimize the role of language, and by replacing linguistic forms of tasks used in prior studies with tasks constructed with only perceptual elements. We showed participants displays containing two vertically arranged objects, with a brief presentation time that we estimated would allow them just enough time to complete one vertical shift of attention between the objects, either toward the top or the bottom object. We then tested for response time asymmetries between two types of objects across two tasks. We call the object that is attended after a vertical attentional shift the *vertical-shift-object*, and the other object the *non-vertical-shift-object*.

Our control task was an identity judgment: which of these two objects—one seen and one new—did you see in the display? Our experimental tasks were two types of spatial relation judgments. Of greatest interest was the spatial ‘recall’ task, which was designed to mimic the type of representation that might be produced by the attention shift mechanism. In the spatial

recall task, the participant was asked to press a key indicating the relative position of one of the objects. If attentional shifts drive the extraction of relations between objects in an asymmetric way, then the response time for a relational memory task should be relatively faster for the top object after an upward vertical shift, or relatively faster for the bottom object after a downward vertical shift. Such shifts should not affect identity memory, because encoding objects' identity does not require shifts of attention. That is, we predicted an interaction between task (spatial recall vs. identity) and object (vertical-shift vs. non-vertical-shift). In addition to the spatial recall task, we tested a second type of spatial relationship recognition task that should allow a more holistic (template-like) form of recognition. Previous studies have suggested that spatial relational judgments may rely on comparing the current relation to a stored template [51–53]. Thus, instead of showing a single object and asking for its relative position, in the *spatial template* recognition task the participant was shown two arrangements of the critical object, either above or below a blank placeholder, and they indicated which arrangement was shown in the encoding display (see Fig 1). Because the spatial template recognition tasks are more exploratory, we save the results and discussion for the end of Experiment 1.

Experiment 1

We presented participants with pairs of colored circles at three different screen locations and measured their response times for indicating either the identities of the objects or their relative spatial positions. We predicted that vertical shifts of attention should improve the representation for the subsequently attended object for spatial relational memory, but not for identity memory. That is, we expected a response time advantage for the spatial recall task following the direction of attention shift: if attention shift is directed upward, spatial recall would be faster for the top object; if attention is directed downward, spatial recall would be faster for the bottom object.

Our displays placed the objects either vertically centered on the screen (with initial fixation at center; people tended to shift their attention upward), or at the upper screen (initial fixation on the bottom object; people tended to shift their attention upward), or at the lower screen (initial fixation on the top object; people tended to shift their attention downward), see Fig 2. The term *vertical-shift-object* is used to refer to the object that is attended after a vertical attentional shift, and the term *non-vertical-shift-object* is used to refer to the other object. Our overall design was a task (identity, spatial recall, spatial template) X screen location (center screen, upper-screen, lower-screen) X object (vertical-shift-object, non-vertical-shift-object) within-subjects design, and our prediction was that there would be a three-way interaction among the factors.

Method

Participants. Nineteen undergraduate students at Northwestern University in the United States participated in a 45-min session in exchange for course credits or payment.

Ethical Considerations. This project was approved by the Northwestern University Institution Review Board under protocol STU00017701. All participants provided written informed consent. This consent procedure was approved by the IRB.

Stimuli and apparatus. All stimuli were created and displayed using MATLAB with Psychtoolbox [54,55] on an Intel Macintosh running OS x 10.10. The experiment was displayed on a 17inch View-Sonic E70fB CRT monitor with 1024 x 786 resolution and a rate of 75Hz. Participant eye movements were recorded using an Eyelink 1000 (SR Research) sampling at 1000 Hz. Head position and a viewing distance of 60 cm were fixed with a chin rest. Measurements below are reported in pixels, and there were approximately 40 pixels per degree of visual angle.

Each memory encoding display consisted of two circles of different colors from a pool of eight possible colors (i.e., red, green, purple, yellow, blue, orange, pink and brown). For the identity memory task, a foil color was randomly chosen from the remaining six colors for each trial. The RGB values of the colors were approximately perceptually equiluminant, chosen from the '8-class qualitative' ColorBrewer color sets (<http://colorbrewer2.org/>). Each circle was centered at the vertices of a virtual square with the width of 50 pixels and the height of 50 pixels. Circle pairs were vertically arranged with a distance of 260 pixels away from each other's center point. The pair was always centered on the screen and was randomly presented at one of three possible screen locations: upper screens (center of the top circle 260 pixels away from the center of the screen), center screens (center of the top circle 130 pixels away from the center of the screen), lower screens (center of the top circle zero pixel away from the center of the screen).

Procedure. Each trial consisted of four phases: a verbal task encoding phase, a memory encoding phase, a memory testing phase and a verbal task testing phase (see Fig 1). Each trial began with a forced fixation (monitored by the eye tracker) at the center of the screen for 500ms, followed by four randomly chosen consonants for 1s. Participants were instructed to remember the letters at the start of the trial (the verbal task encoding phase) and were asked to identify the letter sets at the end of each trial (the verbal task testing phase).

During the encoding phase, one pair of colored circles randomly appeared at one of the three screen locations for 250ms, followed by a 50ms mask. The design required that most participants make *at most* one vertical saccade, but never two vertical saccades, or even a single vertical saccade followed by a potential (and unmeasurable) pre-saccadic attentional shift [56,57]—either of the latter two possibilities would prevent our design from detecting an asymmetric representation created by a single vertical shift of attention. Pilot experiments showed that having display timings be determined by online saccade triggers in our eye tracker led to participants—implicitly or explicitly—delaying eye movements in order to gain longer presentation times. Fixed presentation times shorter than 250ms led to many trials where participants never made a single saccade. Longer presentation times led to many trials with multiple vertical saccades. Thus, 250ms presentation time was chosen based on these pilot results.

During the testing phase, participants were presented with one of three memory tasks: the identity memory task, the spatial recall memory task or the spatial template recognition memory task. The type of memory task was randomized on a trial-by-trial basis, such that participants were not able to predict the type of the task at the start of each trial. This ruled out any possible differences in participants' strategies at the encoding phase.

The testing displays differed according to the type of task (see Fig 1). In the *identity* task participants were asked to recall the color of one circle within the pair. This test screen depicted one circle from the encoding display, as well as a new circle that was not presented within the current trial. Participants were instructed to indicate which one of the two circles belonged to the pair they previously saw by pressing either a left or right arrow key (a sticker with the drawing of a left arrow was pasted onto the z key on keyboard; another sticker with the drawing of a right arrow was pasted onto the x key). All factors were counterbalanced, such that there were 7 trials for each set of conditions (screen locations (3) x positions of the queried circle at encoding (2) x locations of correct answer (2)) with a total of 84 trials.

The *spatial recall* task required participants to recall the relative spatial location of one circle within the pair. The test screen consisted of one queried circle at the center of the screen with two arrows pointing upward and downward. Participants were instructed to press the up or down arrow key on the keyboard to indicate the relative spatial location of the queried circle in relation to the other circle from the same pair. All factors were counterbalanced, such that

there were 14 trials for each set of conditions (screen locations (3) x positions of the queried circle at encoding (2)) with a total of 84 trials.

The *spatial template recognition* task required participants to recognize which one of two configurations correctly depicted the spatial location of one circle within the pair. The test screen consisted of two configurations arranged on the left and right side of the screen, respectively: one configuration depicted the queried circle and a square (a shape that was easily distinguishable from the test object) denoting the location of the other circle in the same pair; the other configuration was the horizontal mirror reflection the first configuration. Participants were instructed to press the left or right arrow key (the same keys as the identity memory task) on the keyboard to indicate the configuration that correctly depicted the relative spatial location of the queried circle in relation to the other circle from the same pair. All factors were counterbalanced, such that there were 7 trials for each set of conditions (screen locations (3) x positions of the queried circle at encoding (2) x locations of correct answer (2)) with a total of 84 trials.

After finishing the primary tasks, participants saw two sets of letters side-by-side on the screen and were asked to identify which one was identical to the set presented at the start of the trial, by pressing either the left or right arrow key. The different set had one randomly chosen letter changed to a new consonant. Eye movements were calibrated and validated at the beginning of the experiment and were re-calibrated halfway through the experiment. Participants were told to put their left hand on the left-and-right key set, and their right hand on the up-and-down key set. To familiarize participants with the procedures and the arrangement of the response keys, participants performed 12 practice trials before the actual experiment. Participants completed a total of 252 trials with 84 trials for each type of tasks.

Results

Because we seek dissociations between the directions of attentional shifts and performance in the memory tasks, we first analyzed participants' eye movement patterns during the encoding phase, as a reflection of where their attention tended to shift to when stimuli appeared at different screen locations, even if their gaze did not overtly shift on all trials. That is, eye movement analyses were used to identify the "signature" eye movement pattern that captures the majority of trials for each screen location. Response time analyses were based on all trials, not just trials with overt eye movements.

Eye movement analyses. Sample eye movement patterns are depicted in Fig 2. We focused on initial saccades within each circle pair during the encoding period. Participants made at least one saccade on majority of the trials ($M = 59\%$, $SD = 29\%$), more than one saccade on a few trials ($M = 8\%$, $SD = 9\%$) and no saccade on some trials ($M = 41\%$, $SD = 29\%$). Saccades landed and transitioned into a second fixation before the appearance of the mask on many trials ($M = 28\%$, $SD = 17\%$). Thus, the presentation times in the current experiment was effective at preventing participants from completing more than one saccade between objects. We analyzed and compared the percentage of first saccades that were directed toward the top objects (upward eye movements) to those directed toward the bottom objects (downward eye movements). Because objects were 25 pixels in radius, saccades that traveled less than 25 pixels on the vertical dimension for the upper and lower screen locations would indicate that participants simply re-inspected the same object, rather than shifting attention between objects. These saccades were therefore excluded from the analysis. As expected, participants preferred to initiate an upward eye movement when the stimuli appeared on the upper screens ($M = 98\%$, $SD = 3\%$) and on the center screens ($M = 85\%$, $SD = 11\%$) (Because the center screen had 15% of trials that were not consistent with the signature "upward" attention shift,

we had performed an analysis excluding those trials. This analysis is included in the supporting information [S2 File](#), which shows the same pattern of results as the main analysis reported in the response time analyses). In contrast, they preferred to initiate a downward eye movement ($M = 90\%$, $SD = 11\%$), when the stimuli appeared on the lower screens.

The initial fixation at the start of the image presentation (immediately following the fixation screen) was on the bottom objects for the upper screens, top objects for the lower screens, or neither object for the center screens. The average encoding time (within the duration of 250ms) for the initial fixation ($M = 215\text{ms}$, $SD = 20\text{ms}$) at the center of the screen was more than two times the average encoding time for second fixations ($M = 78\text{ms}$, $SD = 31\text{ms}$) (on trials where saccades landed before the image transitioned into the mask). Thus, after the initial fixation on the non-vertical-shift objects (bottom objects for the upper screens, top objects for the lower screens) or neither object (center screens), participants tended to make one saccade toward the vertical-shift objects (top objects for the upper and center screens, bottom objects for the lower screens), followed by a second fixation on the vertical-shift objects for some trials.

The second fixation was largely disrupted by the mask display, as indicated by lower percentage of trials and lower average duration times compared to the first fixation. Because the first fixation was always on the non-vertical-shift-object for the upper and lower screen locations, we further analyzed the first fixation patterns for those trials. The encoding time for initial fixations was 214ms ($SD = 23\text{ms}$) for the upper screen and 220ms ($SD = 21\text{ms}$) for the lower screen; this difference is marginally significant, $t(14) = 2.08$, $p = .057$. Thus, it appears that participants spent slightly more time encoding the non-vertical-shift-object at the lower than upper screens, which may contribute to the lower screen locations exhibiting a bigger effect of the interaction between memory task and attention shift (see below).

Response time analyses. Four participants were excluded: two had lower than 90% overall accuracy on the primary tasks, one had lower than 90% accuracy in the verbal load task, and one had an average response times longer than two standard deviations above the average. Of the remaining fifteen participants, the average accuracy in the verbal load task was 96% ($SD = 2\%$). The average accuracy was 96% ($SD = 2\%$) in the identity task, 97% ($SD = 2\%$) in the spatial recall task, and 91% ($SD = 4\%$) in the spatial template recognition task. Trials with incorrect responses or responses in the 1% slowest percentile across participants (a threshold of 2900ms) were removed from the analysis. The average response time was 731ms ($SD = 151\text{ms}$) in the identity task, 888ms ($SD = 202\text{ms}$) in the spatial recall task, and 1001ms ($SD = 167\text{ms}$) in the spatial template recognition task. There were no indications of potential accuracy/speed tradeoffs that would change any of the conclusions in the response times ANOVA below.

As predicted, a 3 (task) X 2 (object) X 3 (screen location) within-subjects ANOVA revealed a significant three-way interaction $F(4, 56) = 6.73$, $p < .001$, partial Eta squared = .32 ([Fig 3](#), see [S1 File](#) in the supporting information for full analyses). For the center screen—the condition where participants made upward saccades and no initial fixation on either object—there was a significant interaction between task and object, $F(1, 14) = 5.19$, $P = .039$, partial Eta squared = .27 ([Table 1](#)). A planned comparison indicated that response times in the spatial recall task were significantly faster (86ms) for the vertical-shift-objects ($M = 827\text{ms}$, $SD = 202\text{ms}$) compared to the non-vertical-shift-objects ($M = 913\text{ms}$, $SD = 215\text{ms}$), $t(14) = 2.41$, $p = .03$, $d = .62$; in contrast, in the identity task there was no significant difference between the vertical-shift-objects ($M = 739\text{ms}$, $SD = 178\text{ms}$) and the non-vertical-shift-objects ($M = 714\text{ms}$, $SD = 164\text{ms}$), $t(14) = .62$, $p = .55$.

For the upper screen—the condition where participants made upward saccades and initially fixated on the bottom objects—there was a significant interaction between task and object, $F(1, 14) = 4.65$, $p = .048$, partial Eta squared = .25. A planned comparison indicated that there was no significant difference in the spatial recall task between the vertical-shift objects

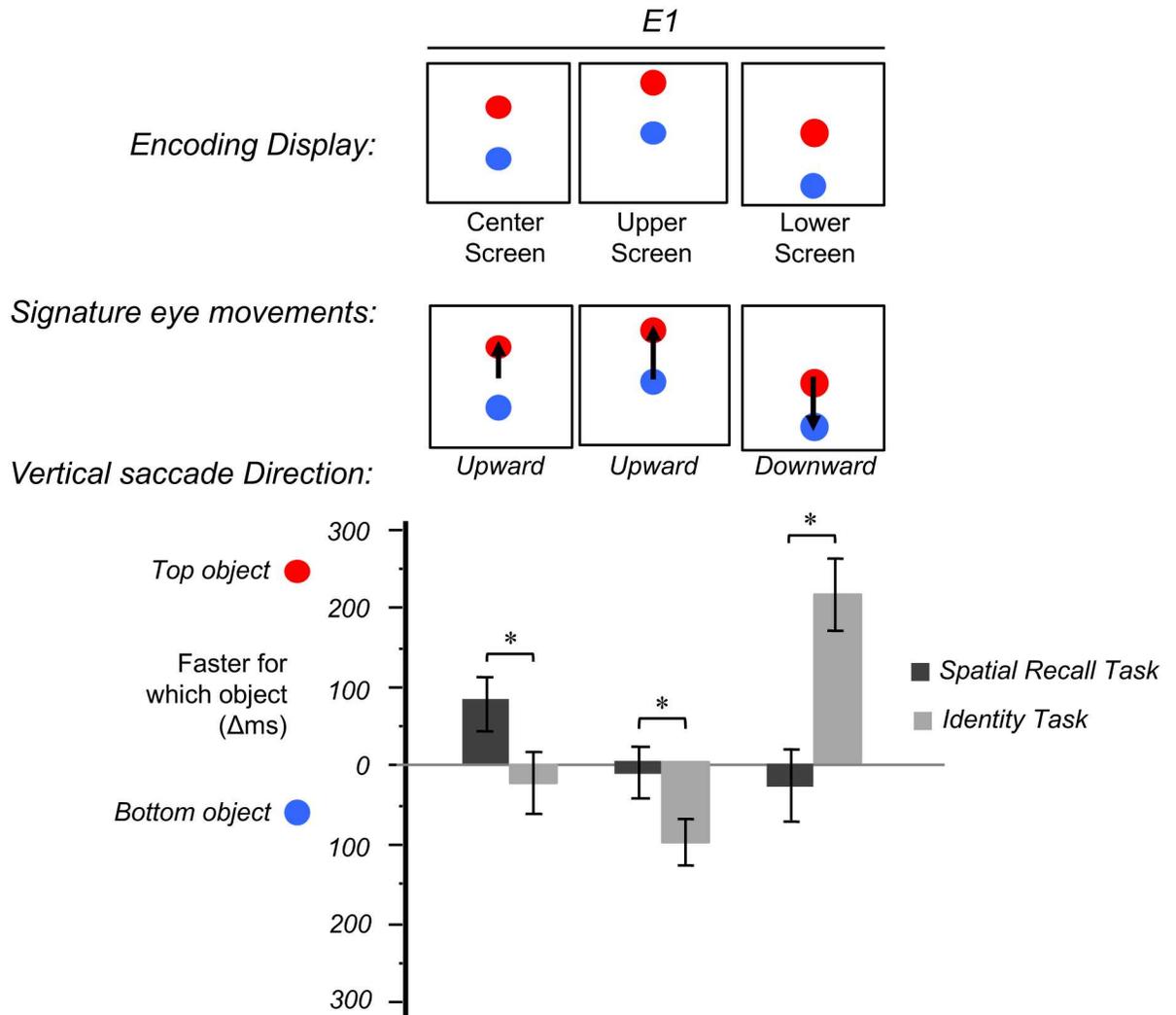


Fig 3. Response times at different encoding displays with corresponding signature eye movements in Experiment 1.

doi:10.1371/journal.pone.0163141.g003

Table 1. Mean reaction time (Standard deviation) for the Vertical-shift-object and Non-vertical-shift-object in the Spatial recall task and the Identity Task across four experiments.

	Screen	Spatial recall task		Identity task		task X object interaction
		Vertical-shift-object	Non-vertical-shift object	Vertical-shift-object	Non-vertical-shift object	
Exp. 1	Center	827ms (202ms)	913ms (215ms)	739ms (178ms)	714ms (164ms)	F (1,14) = 5.19, p = .039
	Upper	898ms (229ms)	883ms (218ms)	789ms (197ms)	683ms (145ms)	F (1,14) = 4.65, p = .048
	Lower	932ms (215ms)	893ms (258ms)	848ms (250ms)	628ms (106ms)	F (1,14) = 15.74, p = .001
Exp.2a	Lower	654ms (71ms)	668ms (117ms)	845ms (197ms)	710ms (133ms)	F (1,11) = 8.17, p = .02
Exp.2b	Lower	751ms (138ms)	818ms (188ms)	778ms (121ms)	720ms (154ms)	F (1,11) = 11.35, p = .006
Exp. 3	Sides	726ms (144ms)	753ms (121ms)	806ms (224ms)	932ms (212ms)	F (1,12) = 4.46, p = .056

Note. Rightmost column depicts F tests for the predicted interaction between task and object for each experiment.

doi:10.1371/journal.pone.0163141.t001

($M = 898\text{ms}$, $SD = 229\text{ms}$) and the non-vertical-shift-objects ($M = 883\text{ms}$, $SD = 218\text{ms}$) objects, $t(14) = .45$, $p = .65$, but response times in the identity task were 106ms faster at the non-vertical-shift-objects ($M = 683\text{ms}$, $SD = 145\text{ms}$) than the vertical-shift-objects ($M = 789\text{ms}$, $SD = 197\text{ms}$), $t(14) = 3.12$, $p = .008$, $d = .88$.

For the lower screen—the condition where participants made downward saccades and initially fixated on the top objects—there was a significant interaction between task and object, $F(1,14) = 15.74$, $p = .001$, partial Eta squared = .52. A planned comparison indicated that there was no significant difference in the spatial recall task between the vertical-shift-objects ($M = 932\text{ms}$, $SD = 215\text{ms}$) and the non-vertical-shift-objects ($M = 893\text{ms}$, $SD = 258\text{ms}$), $t(14) = .64$, $p = .54$, but in the identity task response times were 220ms faster at the non-vertical-shift-objects ($M = 628\text{ms}$, $SD = 106\text{ms}$) than vertical-shift-objects ($M = 848\text{ms}$, $SD = 250\text{ms}$), $t(14) = 4.45$, $p = .001$, $d = 1.6$.

We had no strong predictions for performance in the spatial template recognition task, because we could imagine this processing being similar to spatial recall (shifting attention), or identity (recognizing a relation as a holistic pattern). Interestingly, the results were a mix: similar to those for the spatial recall task for the center and upper screens, but similar to the identity task for the lower screens. For the center screen, there was no significant difference in the spatial template recognition task between the vertical-shift-objects ($M = 922\text{ms}$, $SD = 191\text{ms}$) and the non-vertical-shift-objects ($M = 1014\text{ms}$, $SD = 210\text{ms}$), $t(14) = 1.78$, $p = .1$; however, there was a marginally significant interaction between task (spatial template vs. identity) and object (vertical-shift vs. non-vertical-shift), $F(1,14) = 3.53$, $p = .08$, partial Eta squared = .2. For the upper screen, there was no significant difference in the spatial template recognition task between the vertical-shift-objects ($M = 992\text{ms}$, $SD = 188\text{ms}$) and the non-vertical-shift-objects ($M = 1025\text{ms}$, $SD = 160\text{ms}$), $t(14) = .92$, $p = .37$; however, there was a significant interaction between task (spatial template vs. identity) and object, $F(1,14) = 11.11$, $p = .005$, partial Eta squared = .44. In contrast to the center and upper screens, response times for the spatial template recognition task for the lower screen were significantly faster (168ms) for the non-vertical-shift-objects ($M = 946\text{ms}$, $SD = 219\text{ms}$) than the vertical-shift-objects ($M = 1114\text{ms}$, $SD = 207\text{ms}$), $t(14) = 3.77$, $p < .01$, $d = .99$. However, there was no significant interaction between task (spatial template vs. identity) and object, $F(1,14) = .73$, $p = .41$.

Discussion

Overall, the results confirmed the prediction of a significant interaction between task, screen location and object. Results from the center screen location provided the strongest support for the shift account of relation processing. In that condition, participants spent no significant fixation time on either object at the start of each trial, and showed a dominant strategy of initiating upward eye movements. Spatial recall memory was better for the vertical-shift-objects than the non-vertical-shift-objects. In contrast, identity memory showed no advantage for either objects, consistent with the prediction that extracting object's identity does not require attention shift.

For the upper and lower screen locations, there was also a significant interaction between task (spatial recall vs. identity) and object (vertical-shift-objects and non-vertical-shift-objects). However the pattern of interaction was different from what we expected. In the spatial recall task, rather than an absolute response time advantage for the vertical-shift-objects, we observed a relative advantage when compared to the identity task. Specifically, there was a strong response time advantage for the non-vertical-shift-objects in the identity task, but this advantage was always statistically smaller in the spatial recall task. A possible explanation for this result is that the initial fixation may have benefited the encoding of the non-vertical-shift-

objects [58]. Because the object's identity is more robustly encoded, this may also lead to some response time improvement for that object in the spatial recall task as well. But critically, this spatial recall task advantage for the non-vertical-shift-object was far smaller than that in the identity task: the object's identity is a gateway to its temporary association between that identity and a spatial relationship (top or bottom), but it is not the spatial relationship per se. In other words, although there was a main effect of response time advantage for the non-vertical-shift-objects, there was still a task by object interaction, such that the advantage was smaller or absent for spatial recall task.

We included a spatial template recognition task as an exploratory condition. We did not have a strong prediction for whether performance in that task would be more similar to performance in the spatial recall or identity task. We found the same patterns of results between the spatial template recognition task and the spatial recall task for the center and upper screens, but not for the lower screen (we replicated this result in a separate experiment included in the supporting information [S5 File](#)), which may be tied to differences in attentional resolution or other asymmetries between the vertical hemifields [59,60]. For example, one study found a lower field advantage for processing stationary spatial orienting but an upper field advantage for performing visual search [61].

Experiment 2a

In Experiment 1, stimuli were always directly at, above, or below the initial fixation point, which could lead to artificial strategies that were less related to ecologically relevant decisions about categorical spatial relations—in the real world, relations need to be computed for any set of objects in the two-dimensional visual field. Thus, Experiment 2a (and 2b) sought to replicate the results of Experiment 1 by presenting stimuli at other screen locations (i.e., four corners of a screen). Similar to Experiment 1, we recorded participants' eye movements in Experiment 2a to test if there were differences in the direction of eye movements for stimuli appearing at different screen locations. Experiment 2 omits the spatial template recognition condition, focusing only in the spatial recall and identity tasks. In Experiment 2a, we presented each memory task in separate blocks of trials ([Experiment 2b](#) presented the memory tasks in an mixed and unpredictable order).

Method

Participants. Twelve undergraduate students participated in a 45-min session.

Stimuli and apparatus. All stimuli and apparatus were identical to Experiment 1, with the following exceptions. Each circle was centered at the vertices of a virtual square with the width of 100 pixels and the height of 100 pixels. Each circle pairs were vertically arranged with a distance of 150 pixels away from each other's center point. They were placed at the four corners of the screen 50 pixels away from the edges (see [Fig 4](#)).

Procedure. The procedures were identical to those of Experiment 1, except for the display timing. A longer (compared to [Experiment 1](#)) 400ms display timing was chosen, because it almost always led to a single leftward or rightward saccade, followed sometimes by a second vertical eye movement, and even if not, a presumed vertical attention shift to the second object. There were two types of primary tasks: the identity task and the spatial recall task. The tasks were presented in separate blocks of the trials. The orders of blocks were counterbalanced across participants, such that half participants received the identity task prior to the spatial task, while the other half received the tasks in the reversed order. All factors were counterbalanced. For the spatial recall memory task, there were 16 trials for each set of conditions (screen locations (4) x positions of the queried circle at encoding (2)) with a total of 128 trials. For the identity memory

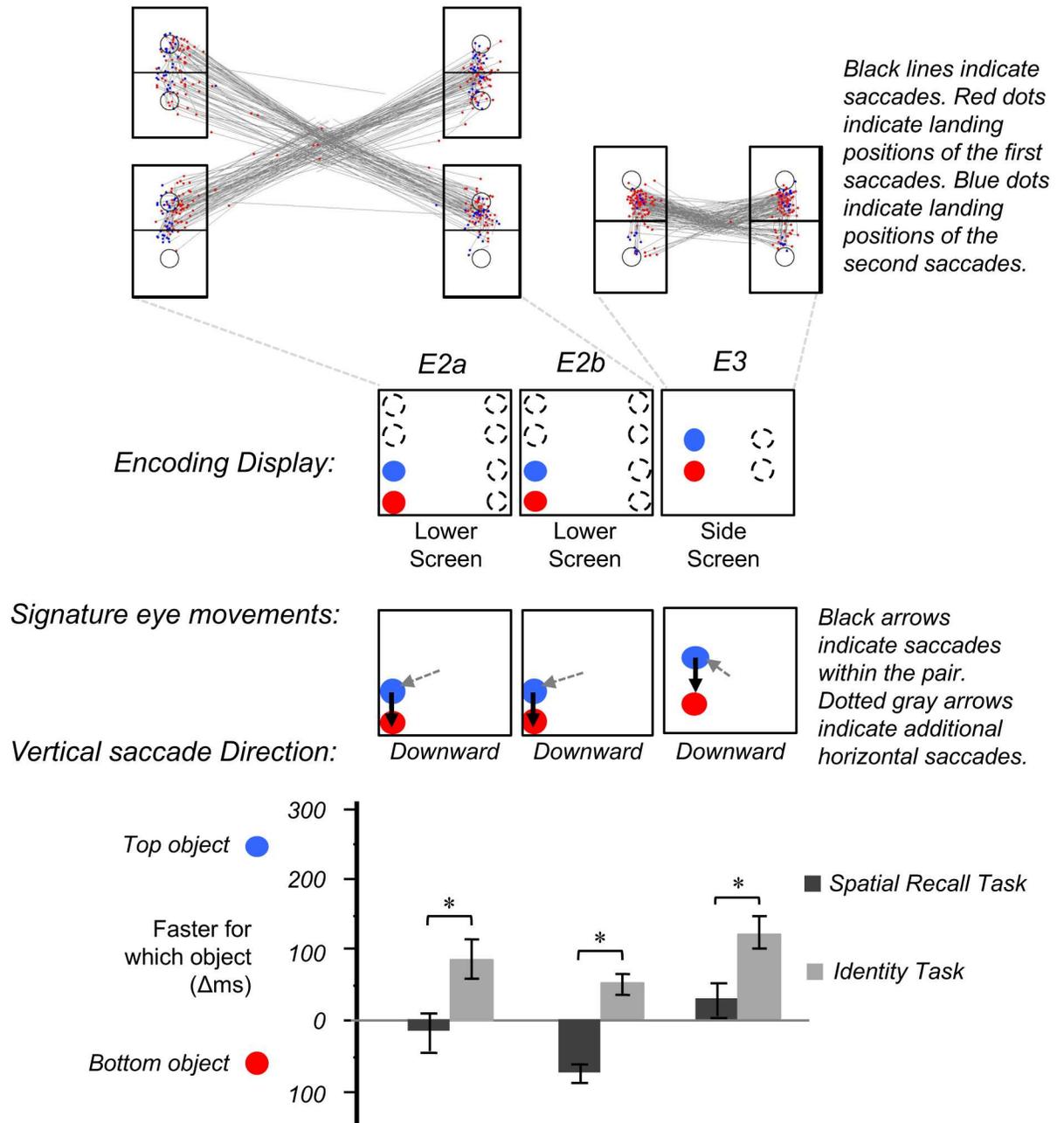


Fig 4. Response times at different encoding displays with corresponding signature eye movements in Experiment 2a, 2b and 3.

doi:10.1371/journal.pone.0163141.g004

task, there were 8 trials for each set of conditions (screen locations (4) x positions of the queried circle at encoding (2) x locations of correct answer (2)) with a total of 128 trials.

Results

As Experiment 1, we first analyzed participants' eye movement patterns during the encoding phase that might reveal differences in the direction of eye movements for stimuli appearing at different screen locations.

Eye movement analyses. Participants made at least one saccade in majority of the trials ($M = 82\%$, $SD = 33\%$), two saccades on some trials ($M = 36\%$, $SD = 23\%$), more than two saccades on very few trials ($M = 3\%$, $SD = 5\%$), and no saccade on some trials ($M = 18\%$, $SD = 33\%$). Before the appearance of the mask, the first saccades landed and transitioned into a fixation on the majority of the trials ($M = 82\%$, $SD = 33\%$), and the second saccades landed and transitioned into a second fixation on many trials ($M = 20\%$, $SD = 17\%$). Thus, after initiating a saccade (a diagonal and non-vertical-saccade) from the center of the screen to the object pair at the corners, most of the time participants made one fixation on one object, sometimes followed by a second between-object-saccade towards the other object and a second fixation on that object. Thus, the presentation times in the current experiment was effective at preventing participants from completing more than one between-object attention shift.

To analyze saccade patterns, we created eight interest areas corresponding to the top and bottom objects within each pair (see Fig 4 for a sample participant). For the first (non-vertical, leftward or rightward) saccades, participants made saccades relatively more often toward the top objects in general. This trend was far more consistent for trials with encoding displays on the lower screens: saccades landed on top objects for 61% ($SD = 18\%$) of trials, on bottom objects for 3% ($SD = 3\%$) of trials, and on neither object for 36% ($SD = 17\%$) of trials. The ratio between the number of saccades that landed on bottom objects to those that landed on top objects was significantly lower than 1 across participants ($M = .04$, $SD = .06$), $t(14) = 55.43$, $p < .0001$. For the upper screens, saccades landed on top objects for 18% ($SD = 21\%$) of trials, on bottom objects for 36% ($SD = 20\%$) of trials, and on neither objects for 46% ($SD = 9\%$) of trials. The ratio between the number of saccades that landed on top objects to those that landed on bottom objects was not significantly different from 1 across participants, ($M = 4.51$, $SD = 11.32$), $t(14) = 1.07$, $p = .31$. This overall pattern is consistent with two previously documented strategies for prioritizing attention and gaze: priorities for objects closer to a center fixation point [12,62,63], and priorities for the top of an object [46,64]. For the lower-screen condition, these priorities converge on a preference for the top object. But for the upper-screen condition, they conflict—the top object is on top, but the bottom object is closer. Indeed, we saw this conflict reflected in the eye movement data. Priorities varied among participants, and across trials with participants. It was additionally unclear how to categorize many of the fixations that landed in the center of the object group as belonging to either the top or bottom object (e.g., Fig 4). We therefore decided to analyze only the data from the trials in the lower-screen display condition. (We provide additional analyses in the supporting information S3 File, focusing only on the upper screen locations. In accordance with the inconsistent patterns of eye movement, we found no benefit for vertical-shift-objects or non-vertical-shift-objects in either task)

Critical to our analysis is the direction of the second vertical saccade between objects. For the lower screens, because first saccades were predominately directed toward the top objects, second saccades that landed on the top objects would indicate that participants simply re-inspected the same object, rather than shifting attention between objects. These saccades were therefore excluded from the analysis. Of the remaining second saccades, the direction was primarily downwards ($M = 87\%$, $SD = 29\%$). Thus, for the lower screen display, after making an initial fixation on the non-vertical-shift-objects (top objects), participants made a downward shift of attention towards the vertical-shift-objects (bottom objects).

Response time analyses. The average accuracy in the verbal load task was 96% ($SD = 3\%$). Average accuracy was 98% ($SD = 2\%$) in the spatial recall memory task and 93% ($SD = 4\%$) in the identity memory task. Trials with incorrect responses or responses in the 1% slowest percentile across participants (a threshold of 3000ms) were removed from the analysis. The average response time was 668ms ($SD = 105$ ms) in the spatial recall memory task, and 782ms ($SD = 140$ ms) in the identity memory task.

A 2 (task) X 2 (object) within-subjects ANOVA revealed a main effect of object, $F(1,11) = 7.67$, partial Eta squared = 0.41, $p = 0.02$ –people in general responded faster to the non-vertical-shift-objects ($M = 689\text{ms}$, $SD = 125\text{ms}$) than the vertical-shift-objects ($M = 750\text{ms}$, $SD = 175\text{ms}$). There was also a main effect of task, $F(1,11) = 14.52$, partial Eta squared = .57, $p = 0.003$. People were faster at the spatial recall memory task ($M = 661\text{ms}$, $SD = 96\text{ms}$) than the identity memory task ($M = 778\text{ms}$, $SD = 178\text{ms}$). As predicted, there was a significant interaction between task and object, $F(1, 11) = 8.17$, $p = .02$, partial Eta squared = .43. A planned comparison indicated that there was no significant difference in the spatial recall task between the non-vertical-shift-objects ($M = 668\text{ms}$, $SD = 117\text{ms}$) and the vertical-shift-objects ($M = 654\text{ms}$, $SD = 71\text{ms}$), $t(11) = .60$, $p = .56$, but in the identity task response times were significantly faster for the non-vertical-shift-objects ($M = 710\text{ms}$, $SD = 133\text{ms}$) than the vertical-shift-objects ($M = 845\text{ms}$, $SD = 197\text{ms}$) objects, $t(11) = 3.21$, $p = .008$.

Discussion

Experiment 2a found that when encoding pairs were presented on the lower screens, participants consistently fixated at the top objects (fixation locations were variable when encoding pairs appeared on the upper screens). Focusing on trials on the lower screens where eye movement were more consistent across participants and trials, we found evidence for the predicted interaction between task (spatial recall task vs. the identity task) and object (vertical-shift-object vs. non-vertical-shift object). Consistent with the results from Experiment 1 upper and lower screen locations, in the spatial recall task, rather than an absolute response time advantage for the vertical-shift-objects, we observed a relative advantage compared to the identity task. That is, there was a strong response time advantage for the non-vertical-shift-objects in the identity task, but this advantage was statistically smaller in the spatial recall task.

Experiment 2b

Experiment 2b replicated the results of 2a while addressing two issues. First, E2a presented the identity and the spatial recall task in separate blocks, which may lead subjects to deploy different strategies for each task. To rule out this possibility, we presented the memory tasks in a mixed and unpredictable order in E2b. Second, because Experiment 2a found consistent eye movement patterns for encoding displays on the lower half but not the upper half of the screen, we presented stimuli at the lower half of the screen for two-thirds of the trials to gain more power for this comparison. We kept one-third of trials on top to induce participants to encode spatial relations between the objects, independent of particular screen locations. We focused our analysis only on the lower screen trials.

Method

Participants. Twelve undergraduate students participated in a 30-min session.

Stimuli and Procedures. All stimuli and procedures were identical to Experiment 2a, except for the following changes: For two-thirds (160) of the total trials (240), encoding pairs were presented at the lower half of the screen. The primary tasks were presented in random order on a trial-by-trial basis. Of the 160 trials, all factors were counterbalanced. For the spatial recall task, there were 20 trials for each set of conditions (screen locations (2) x positions of the queried circle at encoding (2)) with a total of 80 trials. For the identity task, there were 10 trials for each set of conditions (screen locations (2) x positions of the queried circle at encoding (2) x locations of correct answer (2)) with a total of 80 trials. Since the remaining 80 trials at the upper screens were our foil trials, the combinations of factors were presented in a random order without full counterbalancing.

Results

The average accuracy for the verbal load task was 94% ($SD = 2\%$). Average accuracy was 97% ($SD = 2\%$) in the spatial recall memory task, and 95% ($SD = 3\%$) in the identity memory task. Trials with incorrect responses or responses in the 1% slowest percentile across participants (a threshold of 3000ms) were removed from the analysis. The average response time was 784ms ($SD = 165ms$) in the spatial recall memory task, and 749ms ($SD = 135ms$) in the identity memory task.

A 2 (task) X 2 (object) within-subjects ANOVA revealed no main effect of object, $F(1,11) = .02, p = .89$, nor was there a main effect of task, $F(1,11) = .61, p = .45$. However, as predicted, there was a significant interaction between task and object, $F(1,11) = 11.35, p = .006$, partial Eta squared = .51. A planned comparison indicated that participants were 67ms faster at responding to the vertical-shift-objects ($M = 751ms, SD = 138ms$) than the non-vertical-shift-objects ($M = 818ms, SD = 188ms$) at the spatial recall memory task, $t(11) = 1.91, p = .08$, but they were 58ms faster at responding to the non-vertical-shift-objects ($M = 720ms, SD = 154ms$) than the vertical-shift-objects ($M = 778ms, SD = 121ms$) at the identity memory task, $t(11) = 1.85, p = .09$.

Discussion

Experiment 2b replicated the result of Experiment 2a by presenting stimuli predominately at the lower screen locations. Consistent with our hypothesis, we found evidence for the predicted interaction between the spatial recall task and the identity task: identity memory was relatively faster for the non-vertical-shift-objects, but spatial memory was relatively faster for the vertical-shift-objects.

Experiment 3

Experiment 3 replicated the results of Experiment 1, 2a & 2b using 'center-screen' displays more similar to that of Experiment 1, but using the peripheral presentation locations similar to that of Experiment 2. The circle pairs appeared on the left and right side of the screen in vertically centered positions. We again included a relatively small number of corner screen encoding trials (as in [Experiment 2](#)) to encourage participants to encode relations, instead of absolute screen locations, but did not analyze these trials. We recorded eye movements to confirm that, as in Experiment 1, participants prefer to saccade toward the top objects.

Method

Participants. Sixteen undergraduate students participated in a 45-min session.

Stimuli & Procedures. The stimuli and procedures were identical to those in Experiment 2b except for the following changes. During the encoding phase, the vertically aligned circle pairs appeared randomly at either the left or right side of the screen for 75% (192/256) of the trials. To prevent participants using absolute spatial positions in encoding, in the remaining 25% (64/256) of the trials the circle pairs appeared at the four corners of the screen as in Experiment 2b, for a total of six possible presentation locations. Results from the corner trials were not included in the analysis. The circle pairs were horizontally offset from the center of the screen by 150 pixels.

All factors were counterbalanced. For the spatial recall task, there were 24 trials for each set of conditions (screen locations (2) x positions of the queried circle at encoding (2)) with a total of 96 trials. For the identity memory task, there were 12 trials for each set of conditions (screen locations (2) x positions of the queried circle at encoding (2) x locations of correct answer (2))

with a total of 96 trials. Since the remaining 64 trials at the corners were foil trials, the combinations of factors were selected randomly without full counterbalancing.

Results

Eye movement analyses. Participants made at least one saccade in majority of the trials ($M = 92\%$, $SD = 16\%$), two saccades on some trials ($M = 35\%$, $SD = 22\%$), more than two saccades on very few trials ($M = 3\%$, $SD = 3\%$), and no saccade on some trials ($M = 8\%$, $SD = 16\%$). Before the appearance of the mask, the first saccades landed and transitioned into a fixation on the majority of the trials ($M = 89\%$, $SD = 18\%$), and the second saccades landed and transitioned into a second fixation on many trials ($M = 22\%$, $SD = 18\%$). Thus, the presentation times in the current experiment was effective at preventing participants to complete more than one between objects saccade. For the initial saccades, participants preferred to initiate an leftward or rightward eye movement to the top object, for stimuli appearing both at the left ($M = 84\%$ of trials, $SD = 16\%$) and right side ($M = 83\%$ of trials, $SD = 14\%$) of the screen (because 16% of trials were not consistent with the signature “upward” attention shift, we had performed an analysis excluding those trials. This analysis is included in the supporting information [S4 File](#), which shows the same pattern of results as the main analysis reported in the response time analyses). The occasional second saccades were also predominately upward ($M = 76\%$, $SD = 23\%$ for the left side, and $M = 76\%$, $SD = 24\%$ for the right side).

Response time analyses. Three participants were omitted from the analysis due to low accuracy in the verbal load task (75%, 89% and 82%). For the remaining thirteen participants, the average accuracy in the verbal load task was 95% ($SD = 2\%$). Average accuracy was 96% ($SD = 2\%$) in the spatial recall memory task, and 94% ($SD = 3\%$) in the identity memory task. Trials with incorrect responses or responses in the 1% slowest percentile across participants (a threshold of 3000ms) were removed from the analysis. The average response time was 739ms ($SD = 129\text{ms}$) in the spatial recall memory task, and 867ms ($SD = 207\text{ms}$) at the identity memory task.

A 2 (task) X 2 (object) within-subjects ANOVA revealed a main effect of object, $F(1,12) = 12.6$, partial Eta squared = .51, $p = .004$. In general, people were faster at responding to the non-vertical-shift-objects ($M = 766\text{ms}$, $SD = 189\text{ms}$) than the vertical-shift-objects ($M = 842\text{ms}$, $SD = 192\text{ms}$). There was also a main effect of task, $F(1,12) = 5.76$, partial Eta squared = .32, $p = .033$. People were faster at the spatial recall memory task ($M = 739\text{ms}$, $SD = 131\text{ms}$) than the identity memory task ($M = 869\text{ms}$, $SD = 223\text{ms}$). There was again an interaction between task and object, $F(1, 12) = 4.46$, $p = .056$, partial Eta squared = .27. A planned comparison indicated that there was no significant difference between the top objects ($M = 726\text{ms}$, $SD = 144\text{ms}$) and bottom objects ($M = 753\text{ms}$, $SD = 121\text{ms}$) in the spatial recall task, $t(11) = 1.36$, $p = .2$, but response times were significantly faster for the top objects ($M = 806\text{ms}$, $SD = 224\text{ms}$) than the bottom objects ($M = 932\text{ms}$, $SD = 212\text{ms}$) in the identity task, $t(11) = 3.11$, $p = .009$.

Discussion

Experiment 3 used ‘center-screen’ displays more similar to that of Experiment 1, but peripheral presentation locations similar to that of Experiment 2. Once again, we found evidence for the predicted interaction between the spatial recall task and the identity task: identity judgments were relatively faster for the non-vertical-shift-objects, but this advantage was significantly weakened for the spatial recall task. Although the top objects were dominantly fixated, the eye movement that led to that fixation was not vertical—it was primarily toward the left or right of the display, and therefore should not provide a signal that indicates either the top or bottom

object's position in their vertical relation. Consistent with this the shift account, the top object received stronger encoding for the identity judgment, but neither was more strongly encoded for the spatial recall judgment.

General Discussion

We tested the predictions of a low-level mechanistic account of spatial relational processing between two objects—that relations such as 'above' or 'below' are extracted by shifting visual attention upward or downward in space, creating asymmetric relational representations [12,47]. We tested this hypothesis by presenting participants with pairs of colored circles at different screen locations and tested their response times for either the identity or the spatial memory of each object. The objects were presented along the central vertical line of the screen (Experiment 1), at the corners of the screen (Experiments 2a & 2b), and at the left and right sides of the screen (Experiment 3). We also tracked participants' eye movements in Experiments 1, 2a & 3 to reveal how they shifted their attention during encoding.

The strongest support for our central hypothesis comes from the results in the center display in Experiment 1. Specifically, we found that participants predominantly initiated an upward eye movement, but spent no significant fixation time on either object at the start of each trial. This eye movement pattern provides the ideal test for the prediction for how an attention shift between objects can affect the encoding of spatial relations. In this condition, spatial relation recall was better for the vertical-shift-objects than the non-vertical-shift objects. In contrast, identity memory showed no advantage for either object. This result is consistent with the idea that an upward attention shift provides a mental representation of the vertical-shift-object as being on "top". This representation is asymmetric, producing a representation for the relative spatial location of the top vertical-shift-object, but not the bottom object, which did not receive a vertical shift of attention. This preference for directing attention upward when inspecting centrally presented stimuli can be related to spatial asymmetries in the mental representation of time [65], number [66] and mood [10], and maybe due to habitual direction of attention based on culture conventions (e.g., writing direction) [67,68].

Other conditions revealed a second type of eye movement pattern: an initial fixation at the non-vertical-shift-objects, followed by attention shifts to the vertical-shift-objects. For this type of eye movement pattern, attending to an object may enhance the strength of its identity representation, that enhancement should be absent for the object's relational representation, because vertical shift of attention (either overt or covert) toward an object may be required to enhance the strength of its relational encoding. In other words, there should be an interaction between task and object, in which identity memory exhibited a relative benefit for the non-vertical-shift-objects (because of the initial fixation), but this advantage should be absent (and statistically smaller) in spatial recall task. We found this expected performance dissociation between identity and relational memory consistently across Experiment 1 (upper- and lower-screen), Experiment 2a, 2b and Experiment 3. These results are consistent with the attention shift account—there may be a small advantage in spatial recall task for the vertical-shift-objects, but this advantage may be hidden by the main effect of a stronger identity representation for the non-vertical-shift-objects that received more attention. Overall, the pattern of results suggests that attention shift plays a role in the representation of spatial relations.

Although this performance dissociation between identity and relational memory is consistent with the attention shift account, we could imagine how competing accounts could also explain this relative advantage. For example, another account of relation perception specifies that relations are represented by jointly attending objects, but storing their relative locations

and individual features in separate phase cycles of rapidly oscillating activation [35] (but see [21,69,70] for different perspectives). A proponent of that account might claim that the spatial recall task did not enjoy the same most-fixated advantage, because participants did not have time to properly attend to both objects in the display. Similarly, a claim that spatial relations are processed by joint attention through visual indexes [71] might also claim that the second object was not properly indexed. Despite the clear need for further work, we prefer the attentional shift account for several reasons. First, the center display condition of Experiment 1—the one least contaminated by a strong most-fixated object benefit—showed a clear advantage in the spatial recall task for the vertical-shift-objects, which is uniquely consistent with the vertical shift account. Second, the attention shift account provides a low-level mechanism for how the positions of the objects are ‘subtracted’ to compute the positional relation between them—the direction of the vector made by the attentional shift. In contrast, the other accounts provide higher-level descriptions [71]. Third, we find the asymmetries predicted by the attention shift mechanism to be uniquely consistent with the asymmetries found in linguistic descriptions of spatial relations. Past work has argued that the figure/ground or target/referent asymmetries in language might be grounded—or at least compatible with—similar asymmetries in perceptual representations [42,72–75]. The attention shift account of perceptual relation processing could provide an underlying mechanism for that similar representational format. Because the present studies relied on only ‘perceptual’ decisions (not picture-sentence verification) and used a strong verbal suppression dual-task, the present results provide evidence that this asymmetry can arise at a perceptual level, independent of language.

The current study focused on the categorical spatial relation between two objects. Future studies should explore how the shift account of relational processing might relate to other types of spatial coding systems [76,77], including allocentric spatial coding (encoding the location of one object using another object as the reference frame) egocentric spatial coding (using oneself as the reference frame), and geocentric spatial coding (using the environment as the reference frame). Another potentially fertile connection between visual attention and spatial relation encoding is the distinction between categorical and coordinate spatial representation [78]. Several recent studies have found that categorical spatial representation benefits from local attention (e.g., cueing attention to each individual object in a pair), while coordinate spatial representation benefits from global attention (e.g., cueing attention to both objects in a pair simultaneously) [18,19]. Consistent with this view, the current study suggests that isolating individual objects and shifting attention between them plays an important role in categorical spatial representation [12]. While both of these two lines of research emphasize the process, or temporal sequence, of relational processing, given sufficient time, people should be able to extract multiple forms of relational representations (e.g., above and below) utilizing a variety of coding systems [79,80].

In conclusion, the current study provides support for the hypothesis that spatial relations are computed following attention shift between objects [12]. We found a performance dissociation between identity and relation memory: for identity memory, fixating on a given object strongly benefitted response times for that object; in contrast, for relational memory, there was a relative benefit for objects following an upward or downward attention shift. Although further and converging evidence is needed, we suggest that the attention shift account provides a unique low-level explanation for the overall results. The current study suggests that the asymmetric spatial representation reported in previous research [50,72] may exist on the perceptual level, independent of the influence of language and mediated by perceivers’ shifting of attention during online processing [49]. The shift account of relation processing can potentially serve as the common mechanism underlying both linguistic and perceptual representations of spatial relations.

Supporting Information

S1 File. Full analysis of Experiment 1.

(DOCX)

S2 File. Additional analysis of Experiment 1.

(DOCX)

S3 File. Additional analysis of Experiment 2a.

(DOCX)

S4 File. Additional analysis of Experiment 3.

(DOCX)

S5 File. Replicating the Spatial Template Recognition Task (Experiment 4).

(DOCX)

S6 File. Testing Response Keys (Experiment 5).

(DOCX)

Author Contributions

Conceived and designed the experiments: LY SF DU.

Performed the experiments: LY.

Analyzed the data: LY.

Contributed reagents/materials/analysis tools: LY SF.

Wrote the paper: LY SF DU.

References

1. Gentner D. Bootstrapping the mind: analogical processes and symbol systems. *Cogn Sci.* 2010; 34: 752–75. doi: [10.1111/j.1551-6709.2010.01114.x](https://doi.org/10.1111/j.1551-6709.2010.01114.x) PMID: [21564235](https://pubmed.ncbi.nlm.nih.gov/21564235/)
2. Hummel JE, Holyoak KJ. A symbolic-connectionist theory of relational inference and generalization. *Psychol Rev.* 2003; 110: 220–264. doi: [10.1037/0033-295X.110.2.220](https://doi.org/10.1037/0033-295X.110.2.220) PMID: [12747523](https://pubmed.ncbi.nlm.nih.gov/12747523/)
3. Kosslyn SM, Koenig O, Barrett A, Cave CB, et al. Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *J Exp Psychol Hum Percept Perform.* 1989; 15: 723–735. doi: [10.1037/0096-1523.15.4.723](https://doi.org/10.1037/0096-1523.15.4.723) PMID: [2531207](https://pubmed.ncbi.nlm.nih.gov/2531207/)
4. Peli E. Contrast in complex images. *J Opt Soc Am A. Optical Society of America;* 1990; 7: 2032. doi: [10.1364/JOSAA.7.002032](https://doi.org/10.1364/JOSAA.7.002032) PMID: [2231113](https://pubmed.ncbi.nlm.nih.gov/2231113/)
5. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science (80-).* 1981; 211: 453–458. doi: [10.1126/science.7455683](https://doi.org/10.1126/science.7455683) PMID: [7455683](https://pubmed.ncbi.nlm.nih.gov/7455683/)
6. Duffy S, Huttenlocher J, Levine S. It Is All Relative: How Young Children Encode Extent. *J Cogn Dev.* Routledge; 2005; 6: 51–63. doi: [10.1207/s15327647jcd0601_4](https://doi.org/10.1207/s15327647jcd0601_4)
7. Zacks J, Tversky B. Bars and lines: a study of graphic communication. *Mem Cognit.* 1999; 27: 1073–1079. doi: [10.3758/BF03201236](https://doi.org/10.3758/BF03201236) PMID: [10586582](https://pubmed.ncbi.nlm.nih.gov/10586582/)
8. Shah P, Mayer RE, Hegarty M. Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology.* 1999. pp. 690–702. doi: [10.1037/0022-0663.91.4.690](https://doi.org/10.1037/0022-0663.91.4.690)
9. Casasanto D, Boroditsky L. Time in the mind: Using space to think about time. *Cognition.* 2008; 106: 579–593. doi: [10.1016/j.cognition.2007.03.004](https://doi.org/10.1016/j.cognition.2007.03.004) PMID: [17509553](https://pubmed.ncbi.nlm.nih.gov/17509553/)
10. Meier BP, Robinson MD. Why the sunny side is up: association between affect and vertical position. *Psychol Sci.* 2004; 15: 243–7. doi: [10.1111/j.0956-7976.2004.00659.x](https://doi.org/10.1111/j.0956-7976.2004.00659.x) PMID: [15043641](https://pubmed.ncbi.nlm.nih.gov/15043641/)
11. Schubert TW. Your Highness: Vertical Positions as Perceptual Symbols of Power. *J Pers Soc Psychol.* 2005; 89: 1–21. doi: [10.1037/0022-3514.89.1.1](https://doi.org/10.1037/0022-3514.89.1.1) PMID: [16060739](https://pubmed.ncbi.nlm.nih.gov/16060739/)

12. Franconeri SL, Scimeca JM, Roth JC, Helseth S a, Kahn LE. Flexible visual processing of spatial relationships. *Cognition*. Elsevier B.V.; 2012; 122: 210–27. doi: [10.1016/j.cognition.2011.11.002](https://doi.org/10.1016/j.cognition.2011.11.002)
13. Hummel JE. Where view-based theories break down: The role of structure in human shape perception. In: Dietrich E, editor. *Cognitive dynamics: Conceptual and representational change in humans and machines*. Lawrence Erlbaum Associates Publishers. Mahwah, NJ, US; 2000. pp. 157–185.
14. Oliva A, Torralba A. The role of context in object recognition. *Trends Cogn Sci*. 2007; 11: 520–7. doi: [10.1016/j.tics.2007.09.009](https://doi.org/10.1016/j.tics.2007.09.009) PMID: [18024143](https://pubmed.ncbi.nlm.nih.gov/18024143/)
15. Kosslyn SM. *Elements of graph design*. New York: Freeman; 1994.
16. Tanaka K. Columns for Complex Visual Object Features in the Inferotemporal Cortex: Clustering of Cells with Similar but Slightly Different Stimulus Selectivities. *Cereb Cortex*. 2003; 13: 90–99. doi: [10.1093/cercor/13.1.90](https://doi.org/10.1093/cercor/13.1.90) PMID: [12466220](https://pubmed.ncbi.nlm.nih.gov/12466220/)
17. Borst G, Kosslyn SM. Varying the scope of attention alters the encoding of categorical and coordinate spatial relations. *Neuropsychologia*. Elsevier Ltd; 2010; 48: 2769–72. doi: [10.1016/j.neuropsychologia.2010.04.027](https://doi.org/10.1016/j.neuropsychologia.2010.04.027)
18. Laeng B, Okubo M, Saneyoshi A, Michimata C. Processing Spatial Relations With Different Apertures of Attention. *Cogn Sci*. 2011; 35: 297–329. doi: [10.1111/j.1551-6709.2010.01139.x](https://doi.org/10.1111/j.1551-6709.2010.01139.x) PMID: [21429001](https://pubmed.ncbi.nlm.nih.gov/21429001/)
19. van der Ham IJM, Postma A, Laeng B. Lateralized perception: the role of attention in spatial relation processing. *Neurosci Biobehav Rev*. Elsevier Ltd; 2014; 45: 142–8. doi: [10.1016/j.neubiorev.2014.05.006](https://doi.org/10.1016/j.neubiorev.2014.05.006) PMID: [24858009](https://pubmed.ncbi.nlm.nih.gov/24858009/)
20. Hayward WG, Tarr MJ. Spatial language and spatial representation. *Cognition*. 1995; 55: 39–84. doi: [10.1016/0010-0277\(94\)00643-Y](https://doi.org/10.1016/0010-0277(94)00643-Y) PMID: [7758270](https://pubmed.ncbi.nlm.nih.gov/7758270/)
21. Hayward WG. After the viewpoint debate: where next in object recognition? *Trends Cogn Sci*. 2003; 7: 425–427. doi: [10.1016/j.tics.2003.08.004](https://doi.org/10.1016/j.tics.2003.08.004) PMID: [14550482](https://pubmed.ncbi.nlm.nih.gov/14550482/)
22. Hegarty M. The Cognitive Science of Visual-Spatial Displays: Implications for Design. *Top Cogn Sci*. 2011; 3: 446–474. doi: [10.1111/j.1756-8765.2011.01150.x](https://doi.org/10.1111/j.1756-8765.2011.01150.x) PMID: [25164399](https://pubmed.ncbi.nlm.nih.gov/25164399/)
23. Hegarty M, Keehner M, Khooshabeh P, Montello DR. How spatial abilities enhance, and are enhanced by, dental education. *Learn Individ Differ*. 2009; 19: 61–70. doi: [10.1016/j.lindif.2008.04.006](https://doi.org/10.1016/j.lindif.2008.04.006)
24. Stieff M, Bateman RC, Uttal DH. Teaching and Learning with Three-dimensional Representations. In: Gilbert JK, editor. *Visualization in Science Education*. Dordrecht: Springer Netherlands; 2005. pp. 93–120. doi: [10.1007/1-4020-3613-2_7](https://doi.org/10.1007/1-4020-3613-2_7)
25. Landy D, Goldstone RL. How abstract is symbolic thought? *J Exp Psychol Learn Mem Cogn*. 2007; 33: 720–733. doi: [10.1037/0278-7393.33.4.720](https://doi.org/10.1037/0278-7393.33.4.720) PMID: [17576149](https://pubmed.ncbi.nlm.nih.gov/17576149/)
26. Logan GD. Spatial attention and the apprehension of spatial relations. *J Exp Psychol Hum Percept Perform*. 1994; 20: 1015–36. doi: [10.1037/0096-1523.20.5.1015](https://doi.org/10.1037/0096-1523.20.5.1015) PMID: [7964527](https://pubmed.ncbi.nlm.nih.gov/7964527/)
27. Logan GD. Linguistic and Conceptual Control of Visual Spatial Attention. *Cogn Psychol*. 1995; 28: 103–174. doi: [10.1006/cogp.1995.1004](https://doi.org/10.1006/cogp.1995.1004) PMID: [7736720](https://pubmed.ncbi.nlm.nih.gov/7736720/)
28. Palmer J. Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Res*. 1994; 34: 1703–1721. doi: [10.1016/0042-6989\(94\)90128-7](https://doi.org/10.1016/0042-6989(94)90128-7) PMID: [7941377](https://pubmed.ncbi.nlm.nih.gov/7941377/)
29. Wolfe JM. What Can 1 Million Trials Tell Us About Visual Search? *Psychol Sci*. 1998; 9: 33–39. doi: [10.1111/1467-9280.00006](https://doi.org/10.1111/1467-9280.00006) PMID: [25485661](https://pubmed.ncbi.nlm.nih.gov/25485661/)
30. Wolfe JM, Bennett SC. Preattentive Object Files: Shapeless Bundles of Basic Features. *Vision Res*. 1997; 37: 25–43. doi: [10.1016/S0042-6989\(96\)00111-3](https://doi.org/10.1016/S0042-6989(96)00111-3) PMID: [9068829](https://pubmed.ncbi.nlm.nih.gov/9068829/)
31. Hyun J, Woodman GF, Luck SJ. The role of attention in the binding of surface features to locations. *Vis cogn*. 2009; 17: 10–24. doi: [10.1080/13506280802113894](https://doi.org/10.1080/13506280802113894) PMID: [24235876](https://pubmed.ncbi.nlm.nih.gov/24235876/)
32. Luck SJ, Ford M a. On the role of selective attention in visual perception. *Proc Natl Acad Sci U S A*. 1998; 95: 825–830. doi: [10.1073/pnas.95.3.825](https://doi.org/10.1073/pnas.95.3.825) PMID: [9448247](https://pubmed.ncbi.nlm.nih.gov/9448247/)
33. Treisman AM, Gelade G. A feature-integration theory of attention. *Cogn Psychol*. 1980; 12: 97–136. doi: [10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5) PMID: [7351125](https://pubmed.ncbi.nlm.nih.gov/7351125/)
34. Holcombe AO, Linares D, Vaziri-Pashkam M. Perceiving spatial relations via attentional tracking and shifting. *Curr Biol*. Elsevier Ltd; 2011; 21: 1135–9. doi: [10.1016/j.cub.2011.05.031](https://doi.org/10.1016/j.cub.2011.05.031) PMID: [21703857](https://pubmed.ncbi.nlm.nih.gov/21703857/)
35. Hummel JE, Biederman I. Dynamic binding in a neural network for shape recognition. *Psychol Rev*. 1992; 99: 480–517. doi: [10.1037/0033-295X.99.3.480](https://doi.org/10.1037/0033-295X.99.3.480) PMID: [1502274](https://pubmed.ncbi.nlm.nih.gov/1502274/)
36. Franconeri SL. The Nature and Status of Visual Resources. *Oxford handbook of cognitive psychoogy*. 2013. pp. 147–162. doi: [10.1093/oxfordhb/9780195376746.013.0010](https://doi.org/10.1093/oxfordhb/9780195376746.013.0010)
37. Miller GA, Johnson-Laird PN. *Language and perception*. Belknap Press; 1976.
38. Talmy L. How language structures space. In: Pick HL, Acredolo LP, editors. *Spatial orientation*. Boston, MA: Springer US; 1983. pp. 225–282. doi: [10.1007/978-1-4615-9325-6](https://doi.org/10.1007/978-1-4615-9325-6)

39. Landau B, Jackendoff R. Whence and whither in spatial language and spatial cognition? *Behav Brain Sci.* 1993; 16: 255. doi: [10.1017/S0140525X00029927](https://doi.org/10.1017/S0140525X00029927)
40. Gleitman LR, Gleitman H, Miller C, Ostrin R. Similar, and similar concepts. *Cognition.* 1996; 58: 321–376. doi: [10.1016/0010-0277\(95\)00686-9](https://doi.org/10.1016/0010-0277(95)00686-9) PMID: [8871343](https://pubmed.ncbi.nlm.nih.gov/8871343/)
41. Matthews WJ, Dylman AS. The language of magnitude comparison. *J Exp Psychol Gen.* 2014; 143: 510–20. doi: [10.1037/a0034143](https://doi.org/10.1037/a0034143) PMID: [23957280](https://pubmed.ncbi.nlm.nih.gov/23957280/)
42. Clark HH, Chase WG. On the process of comparing sentences against pictures. *Cogn Psychol.* 1972; 3: 472–517. doi: [10.1016/0010-0285\(72\)90019-9](https://doi.org/10.1016/0010-0285(72)90019-9)
43. Clark HH, Chase WG. Perceptual coding strategies in the formation and verification of descriptions. *Mem Cognit.* 1974; 2: 101–111. doi: [10.3758/BF03197499](https://doi.org/10.3758/BF03197499) PMID: [24214706](https://pubmed.ncbi.nlm.nih.gov/24214706/)
44. Taylor HA, Tversky B. Perspective in Spatial Descriptions. *J Mem Lang.* 1996; 35: 371–391. doi: [10.1006/jmla.1996.0021](https://doi.org/10.1006/jmla.1996.0021)
45. Carlson-Radvansky LA, Logan GD. The Influence of Reference Frame Selection on Spatial Template Construction. *J Mem Lang.* 1997; 37: 411–437. doi: [10.1006/jmla.1997.2519](https://doi.org/10.1006/jmla.1997.2519)
46. Xu Y, Franconeri SL. Capacity for Visual Features in Mental Rotation. *Psychol Sci.* 2015; 26: 1241–1251. doi: [10.1177/0956797615585002](https://doi.org/10.1177/0956797615585002) PMID: [26174781](https://pubmed.ncbi.nlm.nih.gov/26174781/)
47. Roth JC, Franconeri SL. Asymmetric coding of categorical spatial relations in both language and vision. *Front Psychol.* 2012; 3: 464. doi: [10.3389/fpsyg.2012.00464](https://doi.org/10.3389/fpsyg.2012.00464) PMID: [23181036](https://pubmed.ncbi.nlm.nih.gov/23181036/)
48. Tomlin RS. Mapping conceptual representations into linguistic representations: The role of attention in grammar. *Language and conceptualization.* 1997. pp. 162–189. doi: [10.1017/cbo9781139086677.007](https://doi.org/10.1017/cbo9781139086677.007)
49. Livins KA, Doumas LAA, Spivey MJ. Tracking Relations : The Effects of Visual Attention on Relational Recognition. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.; 2015.
50. Gleitman LR, January D, Nappa R, Trueswell JC. On the give and take between event apprehension and utterance formulation. *J Mem Lang.* 2007; 57: 544–569. doi: [10.1016/j.jml.2007.01.007](https://doi.org/10.1016/j.jml.2007.01.007) PMID: [18978929](https://pubmed.ncbi.nlm.nih.gov/18978929/)
51. Biederman I, Lescroart M, Hayworth K. Sensitivity to object-centered relations in LOC. *Journal of Vision.* Vision Sciences Society Annual Meeting Abstract; 2007. pp. 1030–1030.
52. Hayworth K. Explicit Encoding of Spatial Relations in the Human Visual System: Evidence from Functional Neuroimaging. 2009.
53. Hayworth K, Lescroart M, Biederman I. Explicit relation coding in the Lateral Occipital Complex. *Journal of Vision.* Vision Sciences Society Annual Meeting Abstract; 2008. pp. 35–35. doi: [10.1167/8.6.35](https://doi.org/10.1167/8.6.35)
54. Brainard DH. The Psychophysics Toolbox. *Spat Vis.* 1997; 10: 433–436. doi: [10.1163/156856897X00357](https://doi.org/10.1163/156856897X00357) PMID: [9176952](https://pubmed.ncbi.nlm.nih.gov/9176952/)
55. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis. Brill;* 1997; 10: 437–442. doi: [10.1163/156856897X00366](https://doi.org/10.1163/156856897X00366)
56. Hoffman JE, Subramaniam B. The role of visual attention in saccadic eye movements. *Percept Psychophys.* 1995; 57: 787–795. doi: [10.3758/BF03206794](https://doi.org/10.3758/BF03206794) PMID: [7651803](https://pubmed.ncbi.nlm.nih.gov/7651803/)
57. Kowler E, Anderson E, Doshier B, Blaser E. The role of attention in the programming of saccades. *Vision Res.* 1995; 35: 1897–1916. doi: [10.1016/0042-6989\(94\)00279-U](https://doi.org/10.1016/0042-6989(94)00279-U) PMID: [7660596](https://pubmed.ncbi.nlm.nih.gov/7660596/)
58. Dickinson CA, Intraub H. Spatial asymmetries in viewing and remembering scenes: Consequences of an attentional bias? 2009; 71: 1251–1262. doi: [10.3758/APP.71.6.1251](https://doi.org/10.3758/APP.71.6.1251) PMID: [19633341](https://pubmed.ncbi.nlm.nih.gov/19633341/)
59. Danckert J, Goodale MA. Ups and downs in the visual control of action. In: Johnson-Frey SH, editor. *Taking action: Cognitive neuroscience perspectives on intentional acts.* MIT Press; 2003. pp. 29–64.
60. Intriligator J, Cavanagh P. The Spatial Resolution of Visual Attention. *Cogn Psychol.* 2001; 43: 171–216. doi: [10.1006/cogp.2001.0755](https://doi.org/10.1006/cogp.2001.0755) PMID: [11689021](https://pubmed.ncbi.nlm.nih.gov/11689021/)
61. Kraft A, Sommer WH, Schmidt S, Brandt S a. Dynamic upper and lower visual field preferences within the human dorsal frontoparietal attention network. *Hum Brain Mapp.* 2011; 32: 1036–1049. doi: [10.1002/hbm.21087](https://doi.org/10.1002/hbm.21087) PMID: [20665723](https://pubmed.ncbi.nlm.nih.gov/20665723/)
62. Wolfe JM, O'Neill P, Bennett SC. Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Percept Psychophys.* 1998; 60: 140–156. doi: [10.3758/BF03211924](https://doi.org/10.3758/BF03211924) PMID: [9503918](https://pubmed.ncbi.nlm.nih.gov/9503918/)
63. Woodman GF, Luck SJ. Serial deployment of attention during visual search. *J Exp Psychol Hum Percept Perform.* 2003; 29: 121–138. doi: [10.1037/0096-1523.29.1.121](https://doi.org/10.1037/0096-1523.29.1.121) PMID: [12669752](https://pubmed.ncbi.nlm.nih.gov/12669752/)
64. Jeerakathil TJ, Kirk A. A representational vertical bias. *Neurology.* 1994; 44: 703–6. doi: [10.1212/WNL.44.4.703](https://doi.org/10.1212/WNL.44.4.703) PMID: [8164830](https://pubmed.ncbi.nlm.nih.gov/8164830/)

65. Boroditsky L, Fuhrman O, McCormick K. Do English and Mandarin speakers think about time differently? *Cognition*. Elsevier B.V.; 2011; 118: 126–132. doi: [10.1016/j.cognition.2010.09.010](https://doi.org/10.1016/j.cognition.2010.09.010) PMID: [21030013](https://pubmed.ncbi.nlm.nih.gov/21030013/)
66. Dehaene S, Bossini S, Giraux P. The mental representation of parity and number magnitude. *J Exp Psychol Gen*. American Psychological Association; 1993; 122: 371–396. doi: [10.1037/0096-3445.122.3.371](https://doi.org/10.1037/0096-3445.122.3.371)
67. Dolscheid S, Shayan S, Majid A, Casasanto D. The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychol Sci*. 2013; 24: 613–21. doi: [10.1177/0956797612457374](https://doi.org/10.1177/0956797612457374) PMID: [23538914](https://pubmed.ncbi.nlm.nih.gov/23538914/)
68. Shaki S, Fischer MH. Reading space into numbers: a cross-linguistic comparison of the SNARC effect. *Cognition*. 2008; 108: 590–9. doi: [10.1016/j.cognition.2008.04.001](https://doi.org/10.1016/j.cognition.2008.04.001) PMID: [18514179](https://pubmed.ncbi.nlm.nih.gov/18514179/)
69. Gentner D, Forbus KD. Computational models of analogy. *Wiley Interdiscip Rev Cogn Sci*. 2011; 2: 266–76. doi: [10.1002/wcs.105](https://doi.org/10.1002/wcs.105) PMID: [26302075](https://pubmed.ncbi.nlm.nih.gov/26302075/)
70. Tarr MJ, Bülthoff HH. Image-based object recognition in man, monkey and machine. *Cognition*. 1998; 67: 1–20. doi: [10.1016/S0010-0277\(98\)00026-2](https://doi.org/10.1016/S0010-0277(98)00026-2) PMID: [9735534](https://pubmed.ncbi.nlm.nih.gov/9735534/)
71. Pylyshyn Z. The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*. 1989; 32: 65–97. doi: [10.1016/0010-0277\(89\)90014-0](https://doi.org/10.1016/0010-0277(89)90014-0) PMID: [2752706](https://pubmed.ncbi.nlm.nih.gov/2752706/)
72. Chase WG, Clark HH. Semantics in the Perception of Verticality. *Br J Psychol*. 1971; 62: 311–326. doi: [10.1111/j.2044-8295.1971.tb02042.x](https://doi.org/10.1111/j.2044-8295.1971.tb02042.x) PMID: [5160435](https://pubmed.ncbi.nlm.nih.gov/5160435/)
73. Just MA, Carpenter P. The Semantics of Locative Information in Pictures and Mental Images. *Br J Psychol*. 1975; 66: 427–441. doi: [10.1111/j.2044-8295.1975.tb01478.x](https://doi.org/10.1111/j.2044-8295.1975.tb01478.x) PMID: [1203643](https://pubmed.ncbi.nlm.nih.gov/1203643/)
74. Seymour PH. Asymmetries in judgments of verticality. *J Exp Psychol*. 1974; 102: 447–455. doi: [10.1037/h0035865](https://doi.org/10.1037/h0035865) PMID: [4815191](https://pubmed.ncbi.nlm.nih.gov/4815191/)
75. Proctor RW, Cho YS. Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychol Bull*. 2006; 132: 416–42. doi: [10.1037/0033-2909.132.3.416](https://doi.org/10.1037/0033-2909.132.3.416) PMID: [16719568](https://pubmed.ncbi.nlm.nih.gov/16719568/)
76. Haun DBM, Rapold CJ, Call J, Janzen G, Levinson SC. Cognitive cladistics and cultural override in Hominid spatial cognition. *Proc Natl Acad Sci U S A*. 2006; 103: 17568–17573. doi: [10.1073/pnas.0607999103](https://doi.org/10.1073/pnas.0607999103) PMID: [17079489](https://pubmed.ncbi.nlm.nih.gov/17079489/)
77. Acredolo LP. Development of spatial orientation in infancy. *Dev Psychol*. 1978; 14: 224–234. doi: [10.1037/0012-1649.14.3.224](https://doi.org/10.1037/0012-1649.14.3.224)
78. Ruotolo F, van Der Ham IJM, Iachini T, Postma A. The relationship between allocentric and egocentric frames of reference and categorical and coordinate spatial information processing. *Q J Exp Psychol*. Psychology Press; 2011; 64: 1138–1156. doi: [10.1080/17470218.2010.539700](https://doi.org/10.1080/17470218.2010.539700)
79. Levelt WJ. Some perceptual limitations on talking about space. *Limits in perception*. VNU Science Press; 1984. pp. 323–358.
80. Burigo M, Coventry KR, Cangelosi A, Lynott D. Spatial language and converseness. *Q J Exp Psychol*. 2016; 1–19. doi: [10.1080/17470218.2015.1124894](https://doi.org/10.1080/17470218.2015.1124894)