

Binding object features to locations: Does the “spatial congruency bias” update with object movement?

Avni N. Bapat¹ · Anna Shafer-Skelton¹ · Colin N. Kupitz¹ · Julie D. Golomb¹

© The Psychonomic Society, Inc. 2017

Abstract One of the fundamental challenges of visual cognition is how our visual systems combine information about an object’s features with its spatial location. A recent phenomenon related to object–location binding, the “spatial congruency bias,” revealed that two objects are more likely to be perceived as having the same identity or features if they appear in the same spatial location, versus if the second object appears in a different location. The spatial congruency bias suggests that irrelevant location information is automatically encoded with and bound to other object properties, biasing perceptual judgments. Here we further explored this new phenomenon and its role in object–location binding by asking what happens when an object moves to a new location: Is the spatial congruency bias sensitive to spatiotemporal contiguity cues, or does it remain linked to the original object location? Across four experiments, we found that the spatial congruency bias remained strongly linked to the original object location. However, under certain circumstances—for instance, when the first object paused and remained visible for a brief time after the movement—the congruency bias was found at both the original location and the updated location. These data suggest that the spatial congruency bias is based more on low-level visual information than on spatiotemporal contiguity cues, and reflects a type of object–location binding that is primarily tied to the original object location and that may only update to the object’s new location if there is time for the features to be re-encoded and rebound following the movement.

Keywords Object–position binding · Spatiotemporal contiguity · Object file · Object recognition

When we look upon a scene, our visual systems typically need to encode multiple different objects. Each of these objects contains a number of features, such as shape, color, and size, which must all be combined for us to recognize what we are seeing. Moreover, we must also process the locations of these objects to successfully act upon them. Object location is not only important for visually guided action, but location has also been proposed to function as a “pointer” or “index” to help individuate objects and successfully solve the “binding problem” (Treisman & Gelade, 1980). For example, when viewing a red pen and a blue mug on your desk, the different features of each object may each be bound to their object’s location (e.g., forming an “object file”; Kahneman, Treisman, & Gibbs, 1992), such that we don’t get confused and incorrectly perceive a blue pen or a red mug.

A number of studies have demonstrated a special role for location in object recognition and a dominance of location information over other types of features across a variety of behavioral tasks (Cave & Pashler, 1995; H. Chen & Wyble, 2015; Z. Chen, 2009; Golomb, Kupitz, & Thiemann, 2014; Leslie, Xu, Tremoulet, & Scholl, 1998; Pertzov & Husain, 2014; Treisman & Gelade, 1980; Tsal & Lavie, 1988, 1993). Of particular interest, a recent line of research has revealed a phenomenon termed the “spatial congruency bias” (Golomb et al., 2014). This bias reveals that people are more likely to judge two sequentially presented objects as having the same identity or features when the objects are presented in the same location, as compared to when the objects are presented in different locations. This is a robust effect that seems to reflect an automatic influence of location information on object perception. Moreover, it is uniquely driven by location: Object

✉ Julie D. Golomb
golomb.9@osu.edu

¹ Department of Psychology, Center for Cognitive & Brain Sciences, Ohio State University, 201 Lazenby Hall, 1827 Neil Avenue, Columbus, OH 43210, USA

location biases judgments of shape, color, orientation, and even facial identity (Golomb et al., 2014; Shafer-Skelton, Kupitz, & Golomb, 2017), yet these features do not bias each other, nor do they bias judgments of location (Golomb et al., 2014). The spatial congruency bias suggests that irrelevant location information is automatically encoded with and bound to other object properties, biasing perceptual judgments. It seems to reveal an underlying assumption of our visual system that stimuli appearing in the same location are likely to be the same object, with the increased tendency to judge two objects as having the same identity presumably resulting from location serving as an indirect link between them, with both objects bound to the same location pointer.

One theoretical account of the spatial congruency bias is that it may be driven by spatiotemporal contiguity; in the real world, objects typically don't disappear and reappear at new locations, so location is generally a reliable cue for "sameness." Spatiotemporal contiguity is known to be a robust cue for object recognition (Burke, 1952; Cox, Meier, Oertelt, & DiCarlo, 2005; Flombaum, Kundey, Santos, & Scholl, 2004; Flombaum & Scholl, 2006; Flombaum, Scholl, & Santos, 2009; Kahneman et al., 1992; Li & DiCarlo, 2008; Mitroff & Alvarez, 2007; Spelke, Kestenbaum, Simons, & Wein, 1995; Wallis & Bülthoff, 2001; Yi et al., 2008). For example, "object files" are thought to rely on spatiotemporal contiguity, with the strength of the spatiotemporal information surpassing the influence of surface feature cues, such as color, size, and shape (Kahneman et al., 1992; Mitroff & Alvarez, 2007; but see Hollingworth & Franconeri, 2009). When stimuli follow a consistent spatiotemporal movement trajectory, we tend to perceive a single object, even if the features have obviously changed; for instance, if subjects view a red stimulus pass behind an occluder, and a green stimulus emerges at the expected temporal and spatial point, subjects tend to perceive a single object that has changed color (Burke, 1952; Flombaum & Scholl, 2006). Infants exhibit a similar reliance on spatiotemporal information (Flombaum et al., 2009; Spelke et al., 1995), as do nonhuman primates—for example, monkeys may behave as if a kiwi fruit has transformed into a lemon when one replaces the other at the same location (Flombaum et al., 2004). Spatiotemporal contiguity also modulates the neural representations of object identity (Li & DiCarlo, 2008; Yi et al., 2008). Moreover, artificially altering spatiotemporal regularities can have substantial effects on subsequent object recognition—for example, when participants are trained with repeated exposure to "swapped" objects (Cox et al., 2005) or faces that change identity as the head smoothly rotates (Wallis & Bülthoff, 2001).

In the original spatial congruency bias paradigm (Golomb et al., 2014), stimuli were presented sequentially at either the same location or different locations, with a 1- to 2-s blank delay between presentations. Thus, spatial location biased feature perception even without temporal contiguity. But what if

the object moved to a new location while maintaining spatiotemporal contiguity? If the congruency bias is sensitive to these contiguity cues, we would expect the bias to track with the moving object and update to reflect its new spatial location. But if the congruency bias is based simply on a low-level binding of features to a location, movement poses an interesting challenge: If features are bound to one spatial location, when an object moves, is the object–location binding automatically updated, or does it remain linked to the original spatial location, such that the features would have to be subsequently rebound to the object's new location?

In the present study, we tested four variations of the spatial congruency bias paradigm with spatiotemporally contiguous object movement. In Experiment 1, a stimulus was briefly presented inside a placeholder object; the placeholder then smoothly moved to a new location, after which a second stimulus appeared at the final placeholder location, at the original location, or at a control location. In Experiment 2, the stimulus itself moved, rather than a placeholder. Experiments 3 and 4 included manipulations of other factors, such as timing and occlusion during movement, respectively. Across all four studies, we found a strong congruency bias at the initial location, which sometimes—but not always—was partially updated to the end of the movement path.

General method

Subjects

The subjects for these experiments were recruited from the Ohio State University. They ranged from 18 to 35 years of age, with normal or corrected-to-normal vision. A sample size of $N = 16$ for each experiment was chosen on the basis of a power analysis of the original spatial congruency bias experiment reported in Golomb et al. (2014), which had a Cohen's $d = 1.01$ and statistical power ($1 - \beta$) of .96. (One experiment included 17 subjects because we overscheduled an extra subject.) Informed consent was obtained for all subjects, and the study protocols were approved by the Ohio State University Behavioral and Social Sciences Institutional Review Board. All subjects were compensated with a small monetary sum or course credit.

Experimental setup

Stimuli were generated using the Psychophysics Toolbox extension (Brainard 1997) for Matlab (Mathworks) and presented on a 21-in. flatscreen CRT monitor. Subjects were seated at a chinrest 60 cm from the monitor. The monitor was color calibrated with a Minolta CS-100 colorimeter.

Eye-tracking

Eye position was monitored using an EyeLink 1000 eye-tracking system recording pupil and corneal reflection position. Fixation was monitored for all experiments. If at any point the subject’s fixation deviated greater than 2°, the trial was aborted and repeated later in the block.

Stimuli

The stimuli were the same as those in Golomb, Kupitz, and Thiemann (2014), modified from the Tarr stimulus set (stimulus images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University; www.tarrlab.org). Stimuli were drawn from ten families of shape morphs created using FantaMorph software (Abrosoft; www.abrosoft.com/). Each of these ten families contained 20 individual exemplar objects (5% morph difference between each image). Within a family, the “body” of the object always remained constant, but the “appendages” could vary in shape, length, or relative location. Stimulus orientation was never varied. Stimuli were presented on a black background and sized 5° × 5° and centered at 7° eccentricity.

Analyses

The spatial congruency bias was calculated as in Golomb et al. (2014), using the signal detection theory¹ formula below, where “hits” are defined as the subject responding “same” when the two objects in fact had the same identity, and “false alarms” were “same” responses when the objects had different identities:

$$\text{Bias(criterion)} = -\frac{z(\text{hit rate}) + z(\text{false alarm rate})}{2}$$

The spatial congruency bias was calculated separately for each subject and location condition and was submitted to random-effects analyses (planned two-tailed *t* tests). Effect sizes were calculated using Cohen’s *d*. Trials on which subjects failed to respond, or responded with reaction times (RTs) beyond 2.5 *SDs* from the subject’s mean RT were excluded (less than 3% of trials for each experiment). Subjects who had an overall task accuracy of less than 55% (indicating noncompliance or inability to perform the task; the criterion was set in advance, consistent with Finlayson & Golomb, 2016; Shafer-Skelton et al., 2017) were excluded from the analyses.

¹ Note that although the bias (criterion) measure in signal detection theory is commonly interpreted as a response- or decision-level bias, biases can also be perceptual (Witt, Taylor, Sugovic, & Wixted, 2015), as is thought to be the case with the spatial congruency bias (Golomb et al., 2014; Shafer-Skelton et al., 2017; see the General Discussion).

(Reanalyzing the data without excluding these subjects did not change the patterns or conclusions.)

Additional measures, such as RTs, proportions of “same” responses, and *d*-prime are reported in Table 1.

Experiment 1: Placeholder movement

In Experiment 1 (Fig. 1), we modified the original Golomb et al. (2014) spatial congruency bias experiment to test whether the bias updated with object movement. As in the original

Table 1 Results from all experiments

		Start	End	Control 1	Control 2
Experiment 1					
<i>p</i> (“same”)	Same identity	.86	.60	.76	
	Different identity	.50	.31	.39	
Reaction time	Same identity	0.746	0.772	0.806	
	Different identity	0.829	0.782	0.828	
Bias		−0.61	0.14	−0.23	
D-prime		1.20	0.80	1.09	
Experiment 2					
<i>p</i> (“same”)	Same identity	.91	.67	.73	.76
	Different identity	.41	.26	.35	.35
Reaction time	Same identity	0.763	0.815	0.796	0.798
	Different identity	0.804	0.826	0.822	0.818
Bias		−0.58	0.12	−0.12	−0.18
D-prime		1.65	1.22	1.09	1.19
Experiment 3					
<i>p</i> (“same”)	Same identity	.74	.77	.71	.70
	Different identity	.41	.41	.36	.39
Reaction time	Same identity	0.738	0.741	0.756	0.757
	Different identity	0.784	0.768	0.771	0.773
Bias		−0.23	−0.27	−0.11	−0.13
D-prime		0.95	1.04	0.98	0.83
Experiment 4					
<i>p</i> (“same”)	Same identity	.79	.74	.71	.72
	Different identity	.54	.53	.48	.51
Reaction time	Same identity	0.425	0.439	0.436	0.426
	Different identity	0.466	0.450	0.436	0.459
Bias		−0.49	−0.38	−0.27	−0.33
D-prime		0.75	0.63	0.64	0.59

Note that the reaction times in Experiment 4 were measured from the time Object 2 was fully unoccluded. However, subjects could have begun accumulating information during the 360-ms partial-occlusion period as the object re-emerged from behind the occluder—hence, the faster reaction times here than in the other experiments

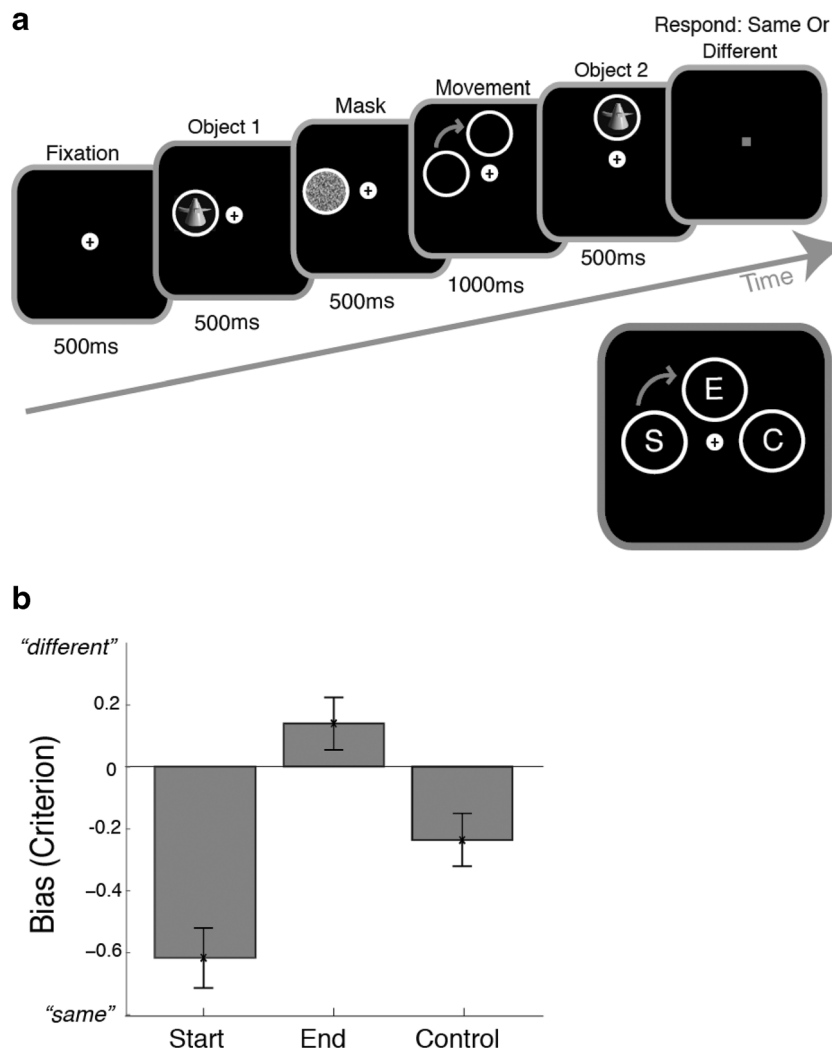


Fig. 1 Task and results for Experiment 1 (placeholder movement). **a** Trial timing: Object 1 appeared inside a white placeholder for 500 ms at the original (start) location. The object was masked, and then the placeholder moved to a new location (end). Object 2 appeared at either the start (S), end (E), or control (C) location—see the inset. The task was to

respond whether Object 1 and Object 2 were the same or different shapes. **b** Bias (the signal detection theory criterion measure) plotted as a function of location condition. A negative bias is an increased likelihood to judge the objects as having the “same shape.” Error bars indicate standard errors of the means (SEMs), $N = 16$

paradigm, subjects saw two sequentially presented objects separated by a brief delay, and they judged whether the objects had the same or different identities. In the present experiment, Object 1 appeared on the screen inside a circular placeholder, the placeholder moved to a new location during the delay, and then Object 2 appeared in one of three locations: at the start location (where the first object was presented), at the end location (where the object would be expected to appear if spatiotemporal contiguity were assumed), or at a control location.

Method

Subjects A total of 16 subjects (11 females, five males; mean age 20.5 years) participated in this experiment; four additional

subjects completed the study but were excluded for poor task performance (accuracy < 55%).

Task and design Participants initiated each trial by fixating on a central fixation cross. After 500 ms of fixation, Object 1 appeared in one of four peripheral locations around the fixation point. The object was presented for 500 ms, surrounded by a circular placeholder (Fig. 1a). The object was then replaced by a mask for 500 ms. The mask then disappeared, leaving only the placeholder remaining. The placeholder moved 90 deg along a circular path, either clockwise or counterclockwise, to the “end” location; motion was continuous and lasted 1,000 ms. Immediately after movement had ceased, Object 2 appeared for 500 ms.

Object 2 was presented in one of three locations, all at equal visual eccentricities: On 50% of the trials, Object 2 appeared

at the end location, inside the placeholder. On 25% of trials, Object 2 appeared at the “start” location of the movement (i.e., the location where Object 1 had been presented). On the remaining 25% of trials, Object 2 appeared at a “control” location, located 90 deg past the end location along the circular trajectory. When Object 2 appeared at the start or control locations, the placeholder reappeared with the objects at these locations. The end location was both the spatiotemporally contiguous condition and the most probable condition.

Subjects indicated whether they thought the two objects were identical (same shape) or not by making a two-alternative forced choice with a “same” or “different” button press. Location was irrelevant to the task; all subjects were instructed to compare the objects’ identities (shape) only. Participants could respond at any point after Object 2’s onset, and accuracy feedback (a green or red square) was given at the end of each trial. After a 2,000-ms intertrial interval, the next trial began.

The identity of Object 1 was randomly chosen from the set of exemplars on each trial; on 50% of trials, Object 2 was the same exact image as Object 1 (“same identity”), and on the other 50% of trials Object 2 was chosen as a different exemplar from the same morph family (“different identity”). The difference between objects was meant to be subtle, and the morph distance was chosen individually for each subject on the basis of a staircase conducted during a practice block before the main task. An adaptive QUEST procedure (Watson & Pelli, 1983) was used, targeting 75% accuracy. The final staircase value from the adaptive training block was selected as the morph distance for the main task. If performance on a certain block dropped below 65% or above 85%, the morph distance was readjusted prior to the next block.

The six different conditions—Object 2 Identity (same or different) \times Object 2 Location (start, end, or control)—were counterbalanced and presented in a randomized order. Each condition was repeated four times per block (eight times per block for the end location conditions, since they were twice as likely), to create blocks of 32 trials each. Subjects completed between eight and ten blocks total in the allotted time for the session. Eye position was monitored in real time on each trial; if subjects broke fixation at any point during the trial, a large red X would appear in the middle of the screen, and the trial was aborted and repeated later in the run. Before beginning the main task, subjects completed one practice block to orient them to the task and determine the appropriate staircase level for identity differences.

Results

Figure 1b shows the bias (criterion) at the three locations where Object 2 could appear. A negative bias indicates a greater tendency to respond “same identity.” Hit and false alarm rates are included in Table 1, along with RTs, bias, and d -prime.

The spatial congruency bias was greatest (i.e., most negative) when Object 2 appeared at the start location, indicating that subjects were more likely to report the two objects as being the same identity when the second object appeared at the same location as the first object, even though this location was inconsistent with spatiotemporal contiguity (and probability) expectations. As in previous reports, the shift in bias was driven by an increase in both hits (reporting the same identity when the objects were the same) and false alarms (reporting the same identity when the objects were actually different). The bias was significantly greater for the start location than for both the control location [$t(15) = -4.68, p < .001, d = 1.17$] and the end location [$t(15) = -8.07, p < .001, d = 2.02$]. Unexpectedly, we also found a bias in the reverse direction at the end location, which was significantly different from that at the control location [$t(15) = 4.36, p = .001, d = 1.09$], meaning that subjects were significantly *less* likely to report these two objects as the same identity in this condition relative to the control.

Discussion

In Experiment 1 we found a spatial congruency bias similar to what has been previously reported (Golomb et al., 2014), in which two objects presented in the same spatial location were more likely to be judged as having the same identity (shape), even though location was irrelevant to the task. However, here we tested whether the congruency bias would update with object movement or would remain at the original spatial location. Interestingly, despite the end location carrying both spatiotemporal contiguity and probability advantages, the spatial congruency bias remained at the start location, where Object 1 had originally been presented. This suggests that the location where an object first appears carries particular importance for the spatial congruency bias and the type of object–location binding that it may reflect.

One alternative possibility is that subjects ignored the movement and maintained attention at the starting location. However, the (unsped) RT results (Table 1) suggest otherwise. RTs were actually fastest at the end location, and RTs for both the start and end locations were significantly faster than those for the control location [start vs. control: $t(15) = -2.74, p = .015, d = 0.86$; end vs. control: $t(15) = -4.31, p < .001, d = 1.08$], indicating that subjects’ attention was successfully moved to the end location with the movement. It thus seems that although spatial attention updated according to the spatiotemporal contiguity and probability cues, the spatial congruency bias did not.

Finally, whereas we found a standard congruency bias at the start location, the results showed an unexpected reverse bias at the end location (subjects were more likely to deem this object “different” than at the control location). As with the standard congruency bias, the reverse bias was driven by

shifts in both hits and false alarms. This effect seems counter-intuitive, since participants' attention seemed to follow the object's placeholder to the end location, yet they perceived the second object as being more different when it appeared at this location. In other words, both the start and end locations had facilitated RTs relative to the control location, but the congruency biases went in opposite directions. In Experiments 2 and 3, we modified the experimental paradigm (by having the object remain visible during movement and manipulating the presentation times), to see whether this reverse bias would persist, or whether the standard congruency bias would update with movement under these conditions.

Experiment 2: Object movement

In Experiment 1 we used a placeholder to signal movement to the new location. But perhaps this movement cue was not effective enough to create spatiotemporal object continuity. In Experiment 2 we asked, would the spatial congruency bias still remain at the start location if the object itself remained visible for the entire duration of the movement? Furthermore, would the reverse bias still persist?

Method

Subjects A total of 17 subjects (nine females, eight males; mean age 19.8 years) participated in this experiment; two additional subjects completed the study but were excluded for poor task performance (accuracy < 55%).

Task and design In the previous experiment, Object 1 was replaced by a placeholder, and the placeholder moved to the end location. In Experiment 2, Object 1 remained visible—there was no placeholder, and the object itself moved to the end location (Fig. 2a). Otherwise, the timing was analogous to the first experiment: Object 1 appeared for 500 ms at the start location, moved for 1,000 ms, and was then masked upon reaching the end location. Then the second object appeared for 500 ms at either the start point, the end point, or a control location. A second control location was also added so that we could control for both the distance effect found in Golomb et al. (2014) and momentum along the movement trajectory (Fig. 2a). The proportions of trials were divided equally among the four location conditions, and location remained irrelevant to the task (same/different identity judgment). The start and end conditions were each compared to both control conditions (note that we did not treat this as a 2×2 design, since we would have had to ignore either the distance or momentum considerations; rather than prioritize one type of control, we compared all combinations using t tests). All other details were the same as in Experiment 1.

Results

Figure 2b illustrates the bias for each of the four location conditions. There was again a strong spatial congruency bias at the start location, with subjects being significantly more likely to report the two objects as having the same identity when Object 2 appeared at the same start location as Object 1 [vs. the end location: $t(16) = -6.99, p < .001, d = 1.70$; Control 1 location: $t(16) = -6.06, p < .001, d = 1.47$; Control 2 location: $t(16) = -5.24, p < .001, d = 1.27$]. Again, we also found a significant reverse bias at the end location, this time relative to both control locations [$t(16) = 2.32, p = .034, d = 0.56$, and $t(16) = 3.46, p = .003, d = 0.84$, for Control 1 and Control 2, respectively]. There was no difference in bias between the two control locations [$t(16) = -1.25, p = .228, d = 0.30$].

Discussion

In Experiment 2, Object 1 remained visible during the movement rather than being replaced by a placeholder. We also included an additional control location. However, we saw the same pattern of results as in Experiment 1. The spatial congruency bias (greater likelihood to report two objects as the same identity) was restricted to the start location, where the first object had originally been presented, despite the fact that the object itself remained visible throughout the movement. Meanwhile, at the spatiotemporally consistent end location, both experiments found a reverse effect, in which the congruency bias to say “same identity” was *less* than at the control locations.

Experiment 3: Timing manipulation

Experiments 1 and 2 demonstrated that the spatial congruency bias does not automatically update when the first object (or its placeholder) moves to a new location. However, just because it did not update automatically under those circumstances does not mean that it never updates. One possibility is that the spatial congruency bias may in fact update with object movement, but rather than updating immediately, it requires time for the object features to be rebound to the new location. To test this account, in Experiment 3 we modified the timing from Experiment 2. Where in Experiment 2 the first object was masked immediately upon completion of the movement trajectory, in Experiment 3 the first object remained visible at the end location for a brief period of time.

Method

A total of 16 subjects (ten females, six males; mean age 18.8 years) participated in this experiment; one additional subject completed the study but was excluded for poor task performance (accuracy < 55%).

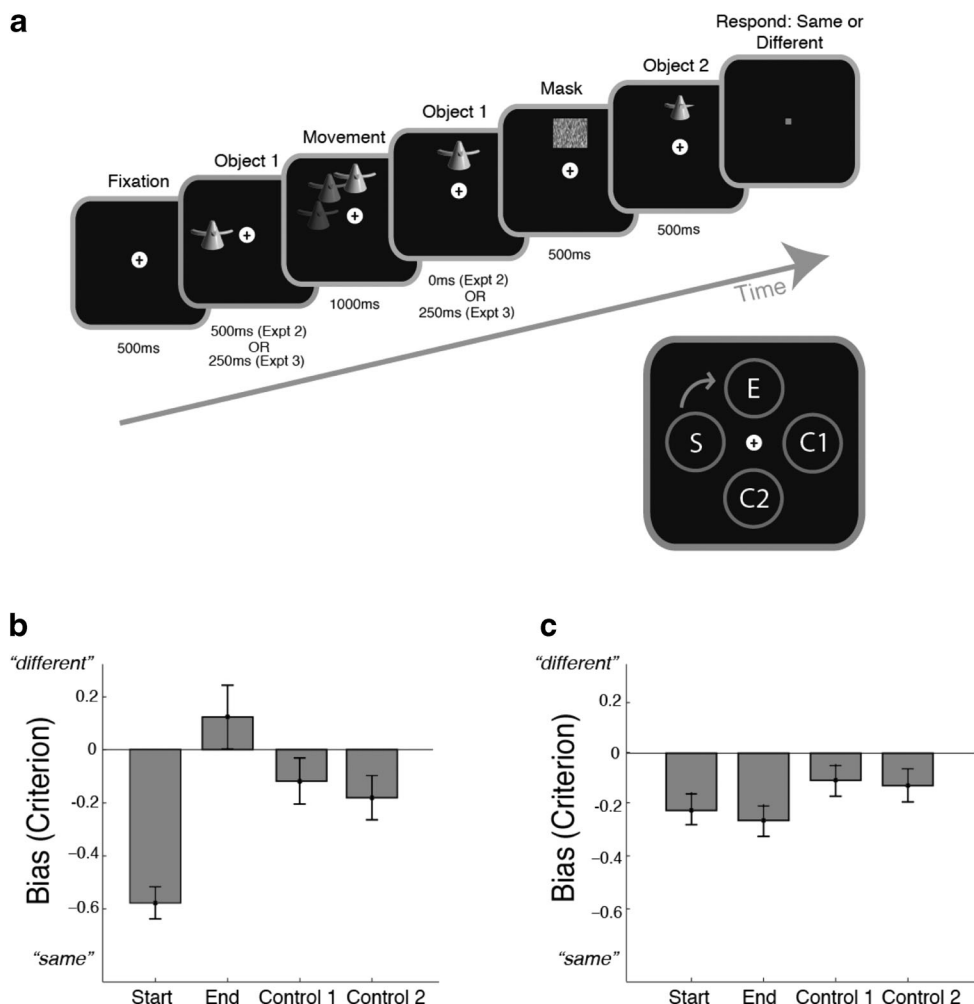


Fig. 2 Task and results for Experiment 2 (object movement) and Experiment 3 (timing manipulation). **a** The trial progression for Experiment 2 was the same as in Experiment 1 (Fig. 1), except that the object itself moved instead of a placeholder. In Experiment 2, Object 1 appeared at the original (start) location for 500 ms and disappeared immediately after the movement to the end location was concluded. In Experiment 3, Object 1 was presented at the start location for 250 ms before the movement, and it remained visible for 250 ms at the end

location. Object 2 appeared at either the start (S), end (E), Control 1 (C1), or Control 2 (C2) location—see the inset. The task was to respond whether Object 1 and Object 2 were the same or different shapes. **b–c** Bias (the signal detection theory criterion measure), plotted as a function of location condition for Experiments 2 (**b**) and 3 (**c**). A negative bias is an increased likelihood to judge the objects as having the “same shape.” Error bars indicate *SEMs*, $N = 17$ (Exp. 2) and $N = 16$ (Exp. 3)

Experiment 3 was identical to Experiment 2 except for the following timing difference: The total object presentation time was equated across experiments, but in Experiment 3 Object 1 was presented before, during, and *after* the movement. Object 1 was presented for 250 ms at the start location, had the same 1,000-ms movement period, and then was visible for 250 ms at the end location before being masked (Fig. 2a).

compared to the previous experiments, as were the effect sizes and statistics [start vs. Control 1: $t(15) = -2.11, p = .05, d = .53$; start vs. Control 2: $t(15) = -1.51, p = .15, d = 0.38$; end vs. Control 1: $t(15) = -2.10, p = .05, d = 0.52$; end vs. Control 2: $t(15) = -2.02, p = .06, d = 0.50$]. There was no significant difference in bias between the start and end locations [$t(15) = 0.56, p = .58, d = 0.14$].

Results

Figure 2c illustrates the bias for each of the four location conditions. The spatial congruency bias now appeared to be split between the start and end locations, with both locations exhibiting a bias to report objects as having the same identity. The magnitude of the spatial congruency bias was reduced as

Discussion

In Experiment 3, we again found a spatial congruency bias at the start location, consistent with the previous experiments. However, the congruency bias was approximately half the magnitude from before, and now seemed to be split between the start and end locations. Whether this means that the bias

was shared between the two locations on each trial, or was present at the start location on some trials and the end location on others, cannot be differentiated here. Regardless, when we reduced the presentation time at the start location and increased it at the end location, Object 1 was now presented for equal amounts of time before and after the movement; under these conditions, the spatial congruency bias seemed to at least partially update to the new location. Note that in previous studies, stimulus presentations of 200 ms were sufficient to evoke a strong congruency bias (Golomb et al., 2014, Exp. 4); thus, even though Object 1's presentation time was split here between the start and end locations, the presentation time of 250 ms at each location should have been sufficient, in principle, to evoke a full bias, though it is possible that the end location bias might have been even larger with additional time after the movement.

However, there is an important caveat for interpreting these results: Because Object 1 was presented both before and after the movement, the updated spatial congruency bias could be explained by a few different accounts. The extra presentation time after the movement could have allowed enough time for the congruency bias to update, but an alternate explanation is that subjects could have simply re-encoded the object at the end location. In other words, it is unclear whether the object–location binding was updated by the spatiotemporal movement or was simply overwritten. Finally, a third possibility is that subjects could technically have performed the task by just paying attention to one location or the other (i.e., sometimes they encoded the object at the start location, but at other times they could have waited until after the movement and just encoded the object then). This ambiguity is why Experiments 1 and 2 were designed in such a way as not to allow the object to be (re)encoded at the end location. Although the results and interpretation of Experiment 3 are more ambiguous than those of the first two experiments, Experiment 3 still makes several important contributions: It shows (1) that the congruency bias can sometimes update, (2) that a sizable component still remains at the start location, and (3) that the reverse bias is eliminated.

Experiment 4: Occluded movement

In the final experiment, we tested whether the spatial congruency bias updates when an object moves behind an occluder. Occluded movement offers the advantage that an object can still be perceived to have spatiotemporal contiguity even when it is no longer visible, thus allowing an opportune time to swap an object's features/identity more subtly without having to rely on an abrupt change or mask. A number of previous studies looking at spatiotemporal contiguity have tested objects moving behind an occluder, demonstrating that when the motion pattern is spatially and temporally consistent, an

object's identity is perceived to remain intact following occlusion (Burke, 1952; Flombaum & Scholl, 2006). Compelling spatiotemporal information can even trump obvious feature differences in the objects before and after the occluder—for example, a kiwi fruit turning into a lemon (Flombaum et al., 2004). Here we tested whether the occluded-motion scenario might provide a more realistic and compelling impression of spatiotemporal contiguity, increasing the likelihood that the object–location binding would update and the congruency bias would transfer to the end location.

Method

A total of 16 subjects (three females, 13 males; mean age 18.8 years) participated in this experiment; eight additional subjects completed the study but were excluded for poor task performance (accuracy < 55%), and one additional subject did not have enough trials remaining after RT trimming.²

In Experiment 4 we used a different layout from the previous versions to accommodate the occlusion (Fig. 3a). Object 1 appeared at one of four possible corner locations around the fixation point; the stimuli were sized $4^\circ \times 4^\circ$ and centered at 9.9° eccentricity. Two occluder bars 4° thick were filled with the random noise mask texture and positioned either horizontally (above and below the fixation point) or vertically (left and right of the fixation point) spanning the entire screen. The occluder bars were repositioned for each trial on the basis of the start location and direction of movement, which was counterbalanced and randomized across trials. The occluders were presented for 500 ms before Object 1 appeared, and they remained visible for the entire trial. Object 1 appeared for 500 ms at the start location and then moved either vertically or horizontally toward the end location. One occluder bar was always positioned near the end of this movement path, such that the object could pass behind the occluder and then reappear at the end location.

The object moved for 830 ms before reaching the occluder, at which point the object disappeared from view. On 25% of trials Object 2 then reemerged from the occluder with the expected spatiotemporal trajectory, and it paused at the end location for 500 ms. (The object was hidden completely behind the occluder for one frame, and then moved for 360 ms

² The majority of the subjects cut for poor accuracy in this experiment also had extremely fast RTs (average RT for the excluded subjects: 190 ms; for the included subjects: 444 ms), indicating that they were not waiting for Object 2 to fully re-emerge but were randomly pressing a button as quickly as possible once it started reappearing. We excluded trials on which the subjects responded within one frame or less of the fully unoccluded object (RTs < 20 ms, a conservative cutoff), and these trials accounted for an average of 16% of trials for the excluded subjects (but <4% for the included subjects). An additional subject had more than 50% of trials excluded for too-fast RTs; although the accuracy for this subject was actually >55% after removal of these trials, there were too few trials remaining to reliably analyze, so this subject was not included in the analyses. An analysis of the data with all subjects included revealed a similar pattern of results to those reported with the subset.

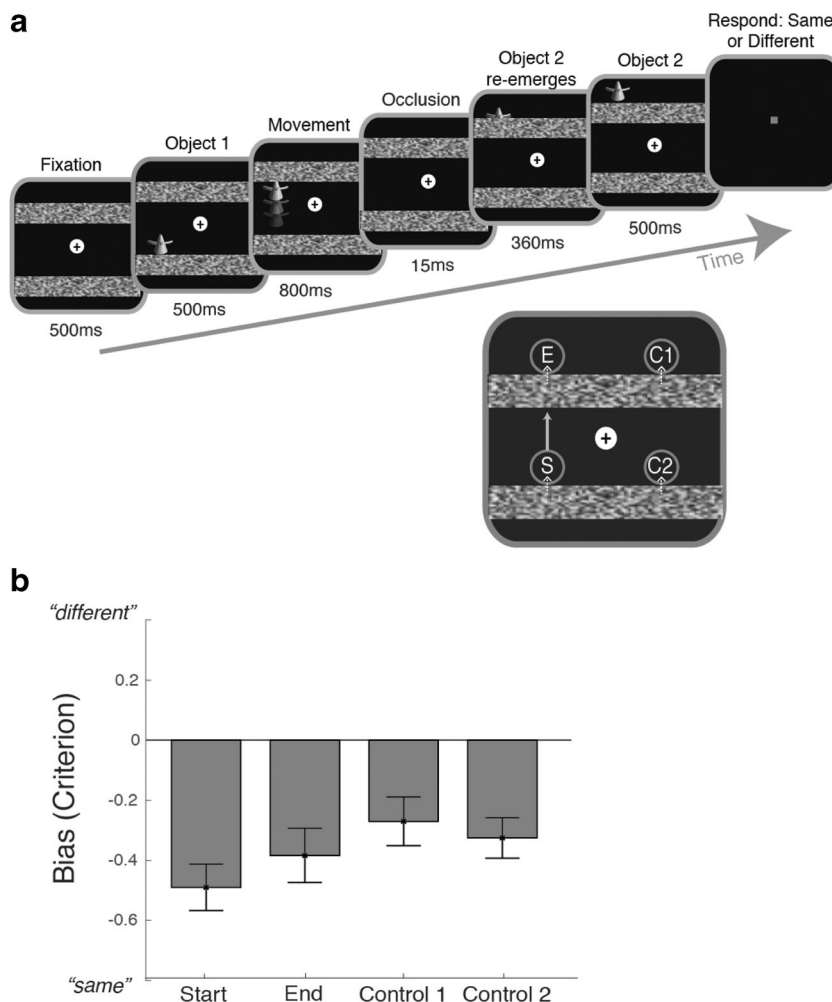


Fig. 3 Task and results for Experiment 4 (occluded movement). **a** Trial timing: Object 1 appeared at the original (start) location for 500 ms and then moved toward and behind an occluder. Object 2 reappeared from behind an occluder and stopped at either the start (S), end (E), Control 1 (C1), or Control 2 (C2) location—see the inset. The task was to respond

whether Object 1 and Object 2 had the same or different shapes. **b** Bias (the signal detection theory criterion measure) plotted as a function of location condition. A negative bias is an increased likelihood to judge the objects as having the “same shape.” Error bars indicate *SEMs*, $N = 16$

until it had reached the fully unoccluded end location position.) On the remaining trials, Object 2 re-emerged from behind a different location on one of the two occluders, and with the same timing moved toward either the original start location or one of two control locations. The four locations were equally likely. Participants responded indicating whether or not they thought the two objects were identical (same shape).

All other details were the same as in Experiments 1–3.

Results

Figure 3b illustrates the biases for each of the four location conditions. The spatial congruency bias was greatest (most negative) at the start location; subjects were significantly more likely to report the two objects as the same identity when the object re-emerged at the start location after occlusion, as compared to both the Control 1 [$t(15) = -4.77, p < .001, d = 1.19$] and Control 2

[$t(15) = -3.64, p = .002, d = 0.91$] locations. The bias at the start location was marginally greater than at the spatiotemporally consistent end location [$t(15) = -2.01, p = .063, d = 0.50$]. The bias at the end location was numerically greater (more negative) than the controls, but it was not significantly different from that at either control location [end vs. Control 1: $t(15) = -1.85, p = .084, d = 0.46$; end vs. Control 2: $t(15) = -1.27, p = .222, d = 0.32$]. There was no significant difference in bias between the two control locations [$t(15) = -0.85, p = .410, d = 0.21$]. The same pattern of results held when all 24 subjects were included, except that the bias at the end location was marginally increased relative to the controls.

Discussion

In Experiment 4 an object appeared, moved across the screen, disappeared behind an occluder, and then re-emerged at either

the spatiotemporally consistent end location, the original start location, or one of two control locations. Despite the strong spatiotemporal expectations at the end location, the spatial congruency bias was again strongest at the start location. The congruency bias at the end location was in the same direction, but was not significantly different from those at the control locations.

General discussion

Here we set out to ask whether the spatial congruency bias (Golomb et al., 2014) is sensitive to spatiotemporal contiguity. The spatial congruency bias is a recently discovered phenomenon demonstrating a robust and dominant effect of an object's spatial location on the perception of its other features and identity. Specifically, two objects appearing in the same location are more likely to be perceived as having the same features or identity. The spatial congruency bias has been proposed to reflect a special role of location information in object recognition. However, the mechanisms and theoretical underpinnings of the congruency bias have yet to be fully uncovered. One critical question is whether the spatial congruency bias is a purely spatial effect reflecting low-level retinotopic input, or whether it is sensitive to more ecologically relevant information about an object's location. Recent studies from our group have shown that the spatial congruency bias remains in retinotopic (not spatiotopic) coordinates after a saccadic eye movement (Shafer-Skelton et al., 2017), and that the congruency bias is driven by 2-D (not 3-D) spatial location information (Finlayson & Golomb, 2016). However, it has long been known that one of the most compelling cues for object "sameness" is spatiotemporal contiguity (Burke, 1952; Flombaum et al., 2009; Hollingworth & Franconeri, 2009; Kahneman et al., 1992; Mitroff & Alvarez, 2007). Given the links to object recognition and the binding problem, an important question is whether the spatial congruency bias is also sensitive to spatiotemporal contiguity, and whether the bias would update with object movement.

In the present study, we tested four variations of the spatial congruency bias paradigm with spatiotemporally contiguous object movement. In Experiment 1, a stimulus was briefly presented inside a placeholder object; the placeholder then smoothly moved to a new location, and a second stimulus appeared at the final placeholder location, at the original location, or at a control location. In Experiment 2, the stimulus itself moved, rather than a placeholder. In both experiments, we found a strong spatial congruency bias at the object's original (start) location. This occurred despite the end location carrying both spatiotemporal contiguity and probability advantages. Interestingly, at the end location there was actually a small effect in the opposite direction in both of these experiments: a reverse bias in which subjects were actually *less*

likely to report the objects as the same, as compared to the control location condition. The meaning of this reverse bias is unclear: We predicted that subjects would either be more likely to report the objects as being the same in this end location than in the control conditions, or that there would be no difference. The fact that we found a reliable reverse bias was unexpected. One possibility is that the spatial congruency bias reflects a more complex mechanism in which facilitation and inhibition interact. Another possibility is that because the objects were either replaced by a placeholder (Exp. 1) or masked at the end of the movement (Exp. 2), this might have interacted with the congruency bias or object file, causing Object 2 to be perceived as "more" different from the original. The fact that the reverse bias was not seen in Experiment 3, when the object remained visible after movement ended, and in Experiment 4, when the object more naturally passed behind an occluder and re-emerged, could support this interpretation, though further study will be needed. Regardless (or perhaps even in spite of this effect), it is notable that the standard-direction spatial congruency bias remained so robust at the start location.

Experiments 3 and 4 included manipulations of timing and occlusion during movement, respectively. In Experiment 3, Object 1 remained visible for an additional period after it had reached the end of the movement path, such that it was presented for equal amounts of time before and after the movement. This led to the elimination of the reverse bias at the end location, with weak evidence for a standard congruency bias at this location as well as at the start location. The congruency bias here was weaker than in the previous experiments, as if it were being split between the start and end locations. However, it is unclear whether the spatial congruency bias actually updated partially to the end location or whether subjects simply re-encoded the object during the delay after the movement.

Finally, Experiment 4 tested a scenario of movement behind an occluder. Object 1 passed behind the occluder near the end of the movement path, which allowed for a more natural transition between Objects 1 and 2. Despite this arguably more compelling sense of spatiotemporal contiguity, the spatial congruency bias again was only reliably present at the original start location.

These results suggest that the location where an object first appears carries particular importance for the spatial congruency bias and the type of object–location binding it may reflect. These results are interesting in light of other work exploring object–location binding, particularly a study by Hollingworth and Rasmussen (2010) looking at object files and visual working memory. Object files are typically probed using the "object-reviewing" paradigm (Kahneman et al., 1992), in which participants tend to display an "object-specific preview benefit": an RT or accuracy advantage when probes subsequently appear on the same object on which they were initially

previewed. This same-object advantage relies heavily on spatiotemporal contiguity (Mitroff & Alvarez, 2007). Hollingworth and Rasmussen used this framework to investigate the role of spatiotemporal contiguity and object files in visual working memory, asking whether object–position binding in visual working memory is linked to the original and/or the updated location after object movement, similar to our question about the spatial congruency bias in the present set of experiments.

In Hollingworth and Rasmussen (2010), four objects were presented simultaneously, in a traditional multi-item working memory paradigm. Each of the four boxes was briefly filled with a color (initial array), the placeholder boxes rotated to new locations (motion), and then the boxes were again filled with colors (test array). The task was to judge whether all of the colors were the same as in the initial array, or whether one was different. The objects in the test array spatially corresponded to the initial positions, the updated positions, or neither. The authors tested RTs and accuracy, finding a performance benefit (bias was untested) for both the original and updated conditions.

Hollingworth and Rasmussen (2010) concluded that two mechanisms of object–position binding were involved in visual working memory: one that updates with motion (object files), and one that is tied to the original location, similar to the object-based and space-based components found for inhibition of return (Tipper, Driver, & Weaver, 1991; Tipper, Weaver, Jerreat, & Burak, 1994). Our finding that the spatial congruency bias is also most strongly tied to the object’s original location, but may sometimes update to the new location, is consistent with these findings.

Interestingly, Hollingworth and Rasmussen’s (2010) latter, motion-insensitive mechanism was interpreted as reflecting a scene-based representation—that is, the features seemed to be bound to the object locations relative to their original configuration in the display (array-centered locations), rather than to their absolute locations. In a recent article, we tested an additional manipulation of the spatial congruency bias—whether it was tied to absolute (spatiotopic) locations or eye-centered (retinotopic) locations following an eye movement (Shafer-Skelton et al., 2017). The spatial congruency bias was linked purely to retinotopic location; even at longer delays after the saccade (more time to update) and for objects of varying complexity (Gabor, objects, and faces), there was no evidence for spatiotopic binding (Shafer-Skelton et al., 2017). Although both paradigms revealed evidence against absolute-position binding, there is a difference between retinotopic (eye-centered) representations (Shafer-Skelton et al., 2017) and configural array-centered representations that survive translation and expansion to different retinotopic positions (Hollingworth & Rasmussen, 2010). Thus, while the spatial congruency bias seems to be primarily associated with the type of object–position binding that is not updated with

motion, it remains unknown whether this nonupdated, retinotopic congruency bias reflects a variation of the configural coding mechanism associated with visual working memory (Hollingworth & Rasmussen, 2010), or whether it reflects a different, third type of binding mechanism. Future research will be needed to test the spatial congruency bias in the presence of multiple-object arrays, scenes, and/or the simultaneous dissociation of retinotopic, spatiotopic, and array-centered reference frames (e.g., Tower-Richardi, Leber, & Golomb, 2016). It is also possible that the spatial congruency bias operates on a different level of perceptual discrimination than the object-specific preview benefit; whereas the congruency bias is only apparent in fine perceptual discrimination tasks and is argued to influence similarity at a perceptual level (Golomb et al., 2014; see the [Discussion](#) below), object-specific preview benefits are typically seen for coarser discriminations (e.g., a set of seven nameable colors in Hollingworth & Rasmussen, 2010) and primarily influence the speed or accuracy of responses.

In sum, the present article adds to a growing body of knowledge characterizing the spatial congruency bias and its relationship to object–location binding. The spatial congruency bias is a robust effect demonstrating that when two objects appear in the same spatial location, they are more likely to be judged to be the same object. The congruency bias is specifically driven by location information—object features such as shape or color do not induce a congruency bias (Golomb et al., 2014). The congruency bias also seems to be more than a simple response-level interference effect—when participants are asked to rate the perceived similarity of two objects on a continuous sliding scale, they systematically rate the objects as being more similar when they appear in the same location versus different locations, and this occurs only when the task is perceptually difficult (Golomb et al., 2014). The spatial congruency bias thus seems to reflect an underlying propensity to use object location as an indicator of “sameness.” When a task is perceptually difficult and two stimuli are not obviously different, our visual systems might rely on the default assumption that if it appears in the same location, it’s probably the same object. If the congruency bias were based purely on some sort of conscious assumption about “sameness,” however, it should be tied to the object’s location in ecologically relevant coordinates. In contrast, the spatial congruency bias appears firmly rooted in the low-level, retinotopic position at which the object was initially encoded. The spatial congruency bias does not seem to automatically update according to spatiotemporal contiguity cues, as we showed in the present experiments, nor does it automatically update to reflect the world-centered spatiotopic location following an eye movement (Shafer-Skelton et al., 2017). Additionally, it has been shown to be sensitive only to the object’s 2-D location on the retina, rather than the object’s 3-D, depth-sensitive position in the world (Finlayson & Golomb, 2016).

The spatial congruency bias thus seems to reflect a low-level, residual effect of the binding of object properties to their original, retinotopic location. As is suggested by the present results, this particular component of binding does not automatically update to the new location when an object moves. It is possible that some additional components of object–location binding do update (Hollingworth & Rasmussen, 2010), or that it is primarily location pointers that automatically update, and that these are dynamically rebound to whatever feature information is present at the new location after the movement is completed (consistent with our findings from Exp. 3). This would be consistent with the idea that spatial tracking of objects preserves limited information about features—for example, in multiple-object tracking, when participants fail to remember the features of objects they are tracking spatially (Horowitz et al., 2007; Pylyshyn, 2004; Saiki, 2003; Scholl, Pylyshyn, & Franconeri, 1999). It could also be consistent with the finding that the *experience* of object continuity is not necessarily tied to the observed object–location binding (Mitroff, Scholl, & Wynn, 2005).

Conclusion

Here we explored a recent phenomenon—the spatial congruency bias—and its role in object–location binding by asking what happens when an object moves to a new location. Across four experiments, we found that the spatial congruency bias remained strongly linked to the original object location. However, under certain circumstances—for example, when the first object paused and remained visible for a brief time after the movement (allowing time to re-encode the object at its new location), the congruency bias was found at both the original location and the updated location. These data suggest that the spatial congruency bias is based more on low-level visual information than on spatiotemporal contiguity cues and reflects a type of object–location binding that is primarily tied to the original object location and may only update to the object’s new location if there is time for the features to rebind following the movement.

Author note This work was supported by research grants from the National Institutes of Health (R01-EY025648) and the Alfred P. Sloan Foundation (BR-2014-098), as well as by support from the Ohio State University Undergraduate Research Office. We thank Samoni Nag and the other members of the Golomb Lab for assistance with data collection and helpful discussion.

References

- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4, 121–138. doi:10.1080/17470215208416611
- Cave, K. R., & Pashler, H. (1995). Visual selection mediated by location: Selecting successive visual objects. *Perception & Psychophysics*, 57, 421–432. doi:10.3758/BF03213068
- Chen, Z. (2009). Not all features are created equal: Processing asymmetries between location and object features. *Vision Research*, 49, 1481–1491. doi:10.1016/j.visres.2009.03.008
- Chen, H., & Wyble, B. (2015). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, 107, 76–85. doi:10.1016/j.visres.2014.11.010
- Cox, D. D., Meier, P., Oertelt, N., & DiCarlo, J. J. (2005). “Breaking” position-invariant object recognition. *Nature Neuroscience*, 8, 1145–1147.
- Finlayson, N. J., & Golomb, J. D. (2016). Feature-location binding in 3D: Feature judgments are biased by 2D location but not position-in-depth. *Vision Research*, 127, 49–56. doi:10.1016/j.visres.2016.07.003
- Flombaum, J. I., Kunder, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in Rhesus Macaques: A study of the tunnel effect. *Psychological Science*, 15, 795–800. doi:10.1111/j.0956-7976.2004.00758.x
- Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 840–853. doi:10.1037/0096-1523.32.4.840
- Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. M. Hood & L. R. Santos (Eds.), *The origins of object knowledge*. New York: Oxford University Press.
- Golomb, J. D., Kupitz, C. N., & Thiemann, C. T. (2014). The influence of object location on identity: A “spatial congruency bias.” *Journal of Experimental Psychology: General*, 143, 2262–2278. doi:10.1037/xge0000017
- Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition*, 113, 150–166. doi:10.1016/j.cognition.2009.08.004
- Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 543–564. doi:10.1037/a0017836
- Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics*, 69, 172–184. doi:10.3758/BF03193740
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219. doi:10.1016/0010-0285(92)90007-O
- Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing “what” and “where” systems. *Trends in Cognitive Sciences*, 2, 10–18. doi:10.1016/S1364-6613(97)01113-3
- Li, N., & DiCarlo, J. J. (2008). Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science*, 321, 1502–1507.
- Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review*, 14, 1199–1204. doi:10.3758/BF03193113
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition*, 96, 67–92. doi:10.1016/j.cognition.2004.03.008
- Pertsov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, & Psychophysics*, 76, 1914–1924. doi:10.3758/s13414-013-0541-y
- Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, 11, 801–822. doi:10.1080/13506280344000518

- Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, 3(1), 6–21. doi:10.1167/3.1.2
- Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? *Investigative Ophthalmology and Visual Science*, 40, S797.
- Shafer-Skelton, A., Kupitz, C. N., & Golomb, J. D. (2017). Object–location binding across a saccade: A retinotopic spatial congruency bias. *Attention, Perception & Psychophysics*, 79, 765–781. doi:10.3758/s13414-016-1263-8
- Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, 13, 113–142. doi:10.1111/j.2044-835X.1995.tb00669.x
- Tipper, S. P., Driver, J., & Weaver, B. (1991). Short report: Object-centred inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, 43, 289–298. doi:10.1080/14640749108400971
- Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 478–499. doi:10.1037/0096-1523.20.3.478
- Tower-Richardi, S. M., Leber, A. B., & Golomb, J. D. (2016). Spatial priming in ecologically relevant reference frames. *Attention, Perception, & Psychophysics*, 78, 114–132. doi:10.3758/s13414-015-1002-6
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. doi:10.1016/0010-0285(80)90005-5
- Tsal, Y., & Lavie, N. (1988). Attending to color and shape: The special role of location in selective visual processing. *Perception & Psychophysics*, 44, 15–21. doi:10.3758/BF03207469
- Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 131–139.
- Wallis, G., & Bühlhoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences*, 98, 4800–4804. doi:10.1073/pnas.071028598
- Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113–120. doi:10.3758/BF03202828
- Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception*, 44, 289–300. doi:10.1068/p7908
- Yi, D.-J., Turk-Browne, N. B., Flombaum, J. I., Kim, M.-S., Scholl, B. J., & Chun, M. M. (2008). Spatiotemporal object continuity in human ventral visual cortex. *Proceedings of the National Academy of Sciences*, 105, 8840–8845. doi:10.1073/pnas.0802525105