



## Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues

Andrew Hollingworth<sup>a,\*</sup>, Steven L. Franconeri<sup>b</sup>

<sup>a</sup> Department of Psychology, The University of Iowa, 11 Seashore Hall E, Iowa City, IA 52242-1407, USA

<sup>b</sup> Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA

### ARTICLE INFO

#### Article history:

Received 16 January 2009

Revised 12 July 2009

Accepted 1 August 2009

#### Keywords:

Object perception

Object persistence

Object correspondence

Occlusion

Dynamic vision

### ABSTRACT

The *correspondence problem* is a classic issue in vision and cognition. Frequent perceptual disruptions, such as saccades and brief occlusion, create gaps in perceptual input. How does the visual system establish correspondence between objects visible before and after the disruption? Current theories hold that object correspondence is established solely on the basis of an object's spatiotemporal properties and that an object's surface feature properties (such as color or shape) are not consulted in correspondence operations. In five experiments, we tested the relative contributions of spatiotemporal and surface feature properties to establishing object correspondence across brief occlusion. Correspondence operations were strongly influenced both by the consistency of an object's spatiotemporal properties across occlusion and by the consistency of an object's surface feature properties across occlusion. These data argue against the claim that spatiotemporal cues dominate the computation of object correspondence. Instead, the visual system consults multiple sources of relevant information to establish continuity across perceptual disruption.

© 2009 Elsevier B.V. All rights reserved.

### 1. Introduction

Although our visual perception of the world seems stable and continuous, the input to vision is in constant flux. Objects move, the observer moves, objects become occluded, and visual input is interrupted by frequent saccades and blinks. Such disruptions cause gaps in the availability of perceptual information. A classic issue in visual cognition is the means by which the visual system establishes correspondence between objects viewed across such disruptions. Consider the following scenario. As you are driving, there are three cars immediately in the road ahead, and you are following one of them to a new restaurant. Your attention is distracted briefly by a nearby pedestrian, and then you switch attention and gaze back to the view ahead. As before, three cars are visible. How does the visual system determine that these are indeed the same three cars visible a moment ago? And, in particular,

how does the visual system establish correspondence between the particular “car to be followed” before and after the disruption? An error in correspondence might mean following the wrong car to the wrong destination. A similar correspondence problem is generated every time one makes an eye movement: how does the visual system determine that an object at one retinal location before a saccade is the same object appearing at a different retinal location after the saccade? Understanding how correspondence problems are solved in vision is central to understanding how we perceive continuity across visual disruptions, such as saccades, and how we track and monitor goal-relevant objects in the service of intelligent behavior.

The visual system could solve correspondence problems using two broad categories of information. First, correspondence could be computed on the basis of *spatiotemporal continuity* across a disruption (Kahneman, Treisman, & Gibbs, 1992; for a review, see Scholl (2007)), with correspondence established if an object's position over time is consistent with the interpretation of a single, persisting

\* Corresponding author. Tel.: +1 319 335 2964; fax: +1 319 335 0191.  
E-mail address: [andrew-hollingworth@uiowa.edu](mailto:andrew-hollingworth@uiowa.edu) (A. Hollingworth).

entity. For example, before being distracted by the pedestrian, the “to be followed” car’s position, speed, and trajectory would be encoded. After the disruption, if there were an object present at the appropriate location given the previous spatiotemporal properties of the tracked car, correspondence would be established, and that object would be treated as a continuation of the tracked car. Second, the correspondence problem could be solved by *surface feature continuity* across a disruption (Hollingworth, Richard, & Luck, 2008; Richard, Luck, & Hollingworth, 2008). When following the car, its perceptual features (white, 4-door, boxy) would be encoded. After a disruption, an object matching the remembered surface features would be treated as a continuation of the previously tracked car. Of course, both spatiotemporal and surface feature information could be consulted and perhaps weighted according to informativeness. If there were many cars in the road ahead moving rapidly through traffic, one might depend more on surface feature properties than on spatiotemporal properties, given difficulty in reliably tracking the positions of the cars. In contrast, if there happened to be multiple white, boxy cars visible, one might depend more on spatiotemporal properties of the cars than on surface feature properties, as surface feature properties would not efficiently distinguish between them.

Although either form of correspondence operation is computationally feasible, researchers have generally concluded that it is the spatiotemporal properties of an object that are used to establish correspondence across the brief disruptions characteristic of dynamic vision (Flombaum, Scholl, & Santos, 2009; Kahneman et al., 1992; Mitroff & Alvarez, 2007; Scholl, 2007). Across longer delays, surface feature information is undoubtedly used to establish object mapping. One’s ability to determine that the dog greeting you in the evening is the same dog you left in the morning clearly depends on memory for the surface feature properties of the dog, a straightforward example of object recognition. In the literature on dynamic object perception, however, the claim is that within a single perceptual event, when one is mapping objects across disruptions such as saccades or brief occlusion, spatiotemporal properties dominate in the computation of object correspondence.

This conclusion has been driven by theory and by empirical evidence. The claim of spatiotemporal dominance follows directly from the most prominent theory of object correspondence, the *object-file* framework of Kahneman et al. (1992). In this view, when an object is present in the visual field, a position marker is assigned to the location occupied by the object (similar to the spatial indexes proposed by Pylyshyn, 2000). If the object moves, the position marker “sticks” to the object, tracking the object’s position over time (i.e., tracking its spatiotemporal properties). Non-spatiotemporal properties of the object (such as color, shape, identity, and so on) can be bound to the position marker, forming the content of the object file.

In the object-file framework, object continuity across change (such as motion) or disruption (such as a saccade or brief occlusion) is established by a *correspondence operation* that consults only the spatiotemporal properties of

the object. If an object’s spatiotemporal properties are consistent across disruption, the same object file assigned to the object before the disruption will be assigned to the object after the disruption, thereby establishing continuity. However, the content of the file, including object surface features and identity, cannot be used to establish correspondence, because these properties can be retrieved from the file only after spatiotemporal correspondence has been established (a process of content *reviewing* that occurs after the correspondence operation). In sum, brief object memory is proposed to be addressed by spatiotemporal properties and not addressed by content (see also Flombaum et al., 2009). Correspondence operations are therefore limited to an analysis of the spatiotemporal properties of objects.

Empirical evidence that object correspondence mechanisms consult only spatiotemporal information comes from the original object-reviewing paradigm developed by Kahneman et al. (1992). In this paradigm, participants see two boxes. Preview letters appear briefly inside each box. The letters are removed, and the empty boxes move to new positions. One test letter appears in a box, and participants name the letter. Naming latency is typically lower when the letter appears in the same box as it had appeared originally versus when it appears in the other box. This *object-specific benefit* has been interpreted as indicating that the box in the test display was treated as the same object that had appeared at a different position at the beginning of the trial (i.e., that object correspondence was established). Specifically, the object-specific benefit suggests that a feature associated with the box at its original location (the preview letter) remains associated with the box at its new location, facilitating letter naming when the test letter matches the letter associated with that box. In the terminology of Kahneman et al., letter naming is facilitated when the test letter matches the content of the object file at that location. The presence or absence of an object-specific benefit across change or disruption therefore provides a measure of the extent to which object correspondence has been established and an object has been treated by the visual system as a single, persisting entity.

The object-specific benefit has been observed consistently when there is spatiotemporal information available to establish object correspondence, such as when participants observe linking motion between two locations of an object (Kahneman et al., 1992). However, when surface feature properties are the only cues to object correspondence, several studies have found no indication of a same object benefit, suggesting that surface feature properties are not consulted in correspondence operations. In Experiment 6 of Kahneman et al. (1992), the letters that appeared in the boxes were drawn in different colors. Two boxes were then moved, in a single step, to new locations equidistant from the original box locations, rendering spatiotemporal properties non-informative. One letter appeared in a box, and the color consistency between the preview and test letter was manipulated. Naming latency was no faster when the preview and test color of the letter were consistent than when they were inconsistent, suggesting that color was not used to establish correspondence between the preview letter and test letter.

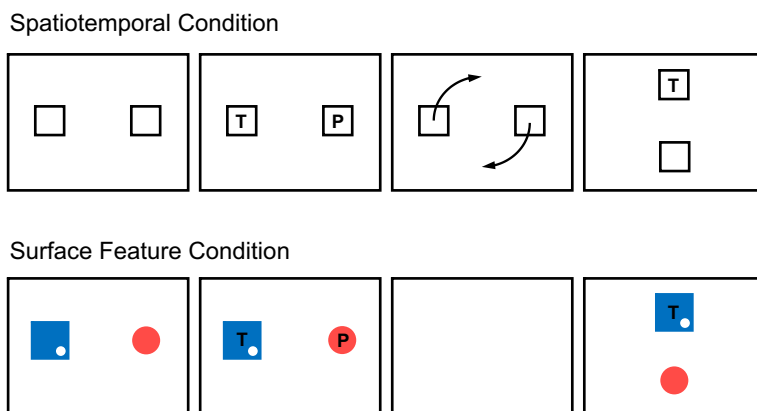


Fig. 1. Illustration of the method used by Mitroff and Alvarez (2007).

Mitroff and Alvarez (2007) modified this paradigm to maximize the possibility of observing correspondence on the basis of a surface feature match. Their method is illustrated in Fig. 1. Participants first saw two objects. In the *spatiotemporal* condition, the objects were identical boxes, as in the basic Kahneman et al. (1992) paradigm. In the *surface feature* condition, the objects differed on a number of salient surface feature attributes, including color, size, shape, and the presence of a hole. Two letters were presented briefly in the two objects, and the objects were then moved to new locations. In the spatiotemporal condition, linking motion between the two positions provided spatiotemporal cues to object correspondence. In the surface feature condition, there was no linking motion; the two boxes were moved in a single step to the new positions across a blank ISI, making spatial correspondence ambiguous. Thus, in the surface feature condition, surface feature information was the only available cue to object correspondence. A test letter was then presented in one of the two objects, and participants reported whether it was or was not one of the two letters appearing at the beginning of the trial.<sup>1</sup>

In the spatiotemporal condition of Mitroff and Alvarez (2007), when the test letter matched one of the preview letters, the test letter could appear either in the same object as it appeared in the preview display (as defined by the motion history of the objects) or in the different object. Similarly, in the surface feature condition, the test letter could appear in the same object in which that letter appeared in the preview display (as defined by the surface feature properties of the objects) or in the different object. An object-specific benefit was observed only in the spatiotemporal condition. That is, only spatiotemporal consistency at test influenced recognition memory perfor-

mance; surface feature consistency produced no reliable effect. Mitroff and Alvarez argued that a surface feature match was not sufficient for the visual system to treat a test object as a continuation of one of the preview objects and that the original object files (containing the preview letter information) were not systematically assigned to test objects on the basis of a surface feature match. Consequently, there was no object-specific effect on recognition memory performance. Given the strength of their surface feature manipulation (the two objects differed on multiple surface feature attributes and were highly distinguishable on the basis of these surface feature differences), the Mitroff and Alvarez data provide the strongest evidence to date that correspondence operations are limited to the spatiotemporal features of an object and do not consult an object's surface features.

The assumption of spatiotemporal dominance is central to understanding how brief forms of object representation are structured in memory and how the visual system is able to maintain coherent representations of dynamically changing environments. In addition, the assumption of spatiotemporal dominance has been central to the translation of the object file framework to other domains of perception and cognition, such as infant object perception (Feigenson & Carey, 2005; Xu, Carey, & Welch, 1999), multiple-object tracking (Horowitz et al., 2007; Pylyshyn, 2004; Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999), and apparent motion (for a review, see Scholl, 2007). We briefly discuss each of these literatures in turn.

In the infant cognition literature, evidence that young infants tend to rely on spatiotemporal and numerical information when interpreting perceptual scenes has been explained as early dependence on an object-file system that does not consult surface feature or identity information (Leslie, Xu, Tremoulet, & Scholl, 1998). However, by one year of age (Xu & Carey, 1996; Xu et al., 1999) or even earlier (Moore, Borton, & Darby, 1978), infants can solve object correspondence on the basis of surface feature properties as well. Although spatiotemporal mechanisms may be primary in their developmental appearance, it does not necessarily follow that they are primary in adult correspondence operations. That is, cognitive processes that

<sup>1</sup> This type of recognition memory task has replaced the original letter naming task of Kahneman et al. (1992) and is now standard in the object-file literature (Kruschke & Fragassi, 1996; Mitroff & Alvarez, 2007; Mitroff, Scholl, & Wynn, 2004; Mitroff, Scholl, & Wynn, 2005). The logic of the recognition memory task is essentially the same as that for the letter naming task. On trials when a test letter is presented within the same object as that letter appeared originally, comparison operations will be relatively efficient, generating an object-specific benefit on elapsed time to respond "same".

arise relatively late in development need not be relegated to a secondary role in adult cognition (language, which develops relatively late, is certainly a central feature of adult, human cognition). Thus, the infant cognition literature does not strongly constrain theorizing about the contribution of spatiotemporal and surface features to correspondence operations in adults.

In work on multiple-object tracking, participants track a set of moving objects among other moving objects (Pylyshyn & Storm, 1988). Object correspondence does indeed appear to be established on the basis of spatiotemporal properties in this paradigm. Participants can successfully track the locations of multiple-objects without necessarily keeping track of their surface feature properties (Horowitz et al., 2007; Pylyshyn, 2004; Scholl, Pylyshyn, & Franconeri, 1999). However, dependence on spatiotemporal properties is not particularly surprising in a task that requires participants to explicitly track the locations of objects over time.

In the literature on apparent motion, motion correspondence can be perceived between two objects that differ in surface feature properties (Kolers, 1972), and ambiguous motion tends to be resolved on the basis of spatiotemporal features rather than on the basis of surface features (for a review, see Dawson, 1991), providing some of the strongest evidence in support of the spatiotemporal dominance hypothesis. Further, in the *tunnel effect*, an object passing behind an occluder is often perceived as a single object in motion if its spatiotemporal properties are consistent with that interpretation, even if its surface features change during occlusion (Flombaum et al., 2009). Despite this evidence of spatiotemporal dominance in motion correspondence, Moore, Mordkoff, and Enns (2007; see also Moore and Enns, (2004)) recently found that an abrupt change in the surface features of an object (such as a change in size or color) during a series of apparent motion frames blocked the interpretation of correspondence, disrupting the perception of continuous motion. This occurred despite the presence of spatiotemporal information consistent with object continuity. Thus, under some circumstances, surface feature properties can trump spatiotemporal properties in the perception of motion.

Finally, the object-file framework has been translated to the domain of correspondence across saccadic eye movements (Henderson, 1994; Henderson & Siefert, 2001; Irwin, 1992; Irwin & Andrews, 1996; Irwin & Gordon, 1998; Zelinsky & Loschky, 2005), but it has been done so without a strong commitment to the assumption of spatiotemporal dominance (Currie, McConkie, Carlson-Radvansky, & Irwin, 2000; Gordon, Vollmer, & Frankl, 2008; Hollingworth et al., 2008; Richard et al., 2008). Recently, Richard et al. tested the relative contributions of spatiotemporal and surface feature information to object correspondence across saccades. The weighting of the two sources of information was governed by the demands of saccade target selection. If a saccade target was selected on the basis of its position, an object in the appropriate position after the saccade tended to be treated as the original target object. However, if a saccade target was selected on the basis of its surface features (such as its color), an object with the appropriate surface features after the saccade tended to be treated as the original saccade target.

In summary, the evidence relevant to assessing the spatiotemporal dominance hypothesis is ambiguous. In the original object-reviewing paradigm (Kahneman et al., 1992; Mitroff & Alvarez, 2007) and in the multiple-object tracking literature, spatiotemporal properties appear to dominate correspondence operations, with little or no consultation of surface features. In the apparent motion literature, spatiotemporal properties tend to be weighted more heavily than surface feature properties, but this can be reversed if there is a salient change in surface features (Moore, Mordkoff, & Enns, 2007). Finally, in the eye movement literature, there is little evidence of spatiotemporal dominance: Surface and spatiotemporal features are consulted flexibly on the basis of the demands of saccade target selection (Richard et al., 2008).

## 2. The present study

The strongest evidence for spatiotemporal dominance comes from the object-reviewing paradigm that originally generated object-file theory (Kahneman et al., 1992). Any general account of the role of spatiotemporal and surface feature information in object correspondence must address this paradigm as one of the central domains of study. As reviewed above, direct tests of surface feature consistency in the object-reviewing paradigm have found no evidence that surface features play any role in correspondence (Kahneman et al., 1992; Mitroff & Alvarez, 2007). However, two aspects of the design of these studies may have limited their sensitivity to effects of surface feature consistency.

First, the spatiotemporal and surface feature conditions were not equated for discontinuity on the two dimensions. Consider the Mitroff and Alvarez method illustrated in Fig. 1. In the spatiotemporal condition, there was no discontinuity in the surface feature properties of the objects; the two boxes retained their original surface features throughout the trial. However, in the surface feature condition, there was a salient discontinuity in spatial information, because the two objects were shifted, in a single step, to new locations, without traversing a path from the original location to the new location. The salient spatiotemporal discontinuity in the surface feature condition could have masked an effect of surface feature consistency. A similar spatiotemporal discontinuity was present in the Kahneman et al. experiment probing color consistency: When the two boxes appeared at new locations, boxes remained visible at the original locations for 87 ms, introducing strong spatiotemporal cues against correspondence. In general, then, to place spatiotemporal and surface feature information on an equal footing, the object-reviewing paradigm would need to be modified so that salient spatiotemporal discontinuity was eliminated from conditions testing the effect of surface feature consistency.

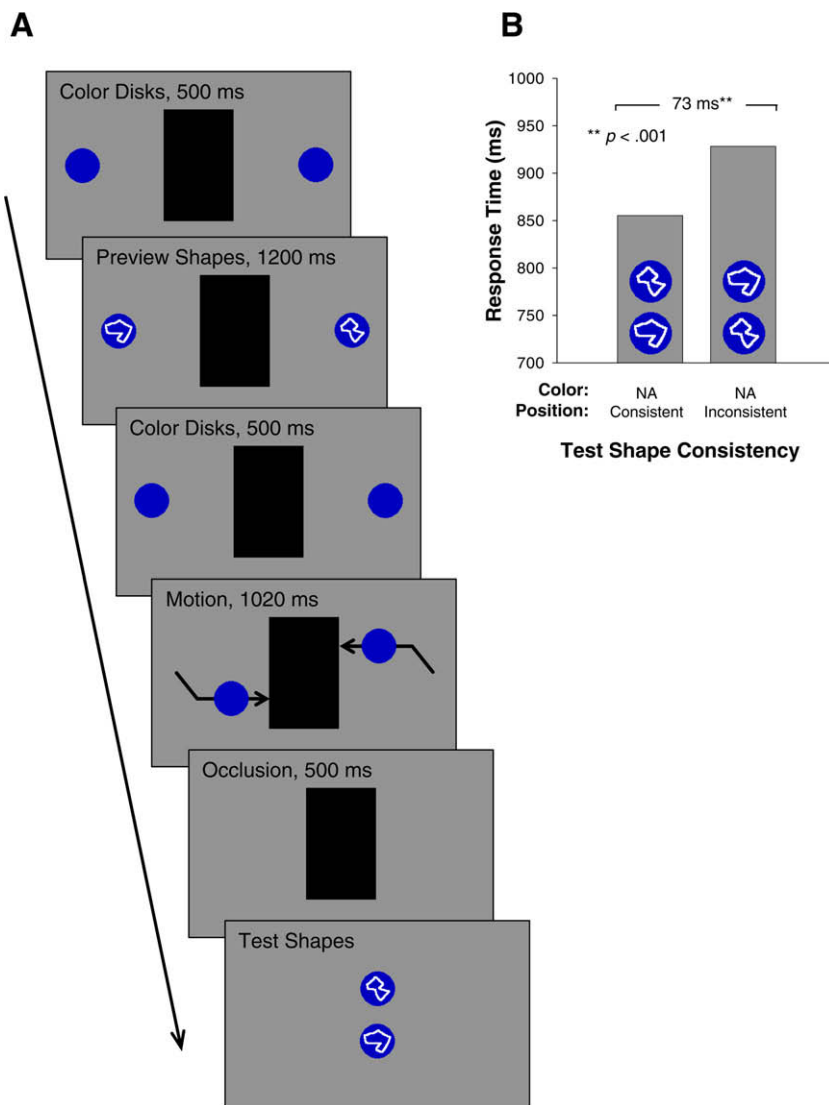
A second limitation of previous experiments testing surface feature correspondence is the use of letters at the test stimuli. Letters certainly can be encoded visually, but they also can be encoded verbally. The absence of a surface feature consistency effect in Kahneman et al. (1992) and Mitroff and Alvarez (2007) might have arisen because the letters and the surface feature properties of the objects in

which they appeared were stored in separate working memory systems (verbal working memory for the letters; VWM for the surface feature properties of the objects) and were not strongly bound together within an object-based, VWM representation (Luck & Vogel, 1997). To provide a stronger test of surface feature correspondence, the test stimuli should require visual memory and should not be easily encodable as verbal labels.

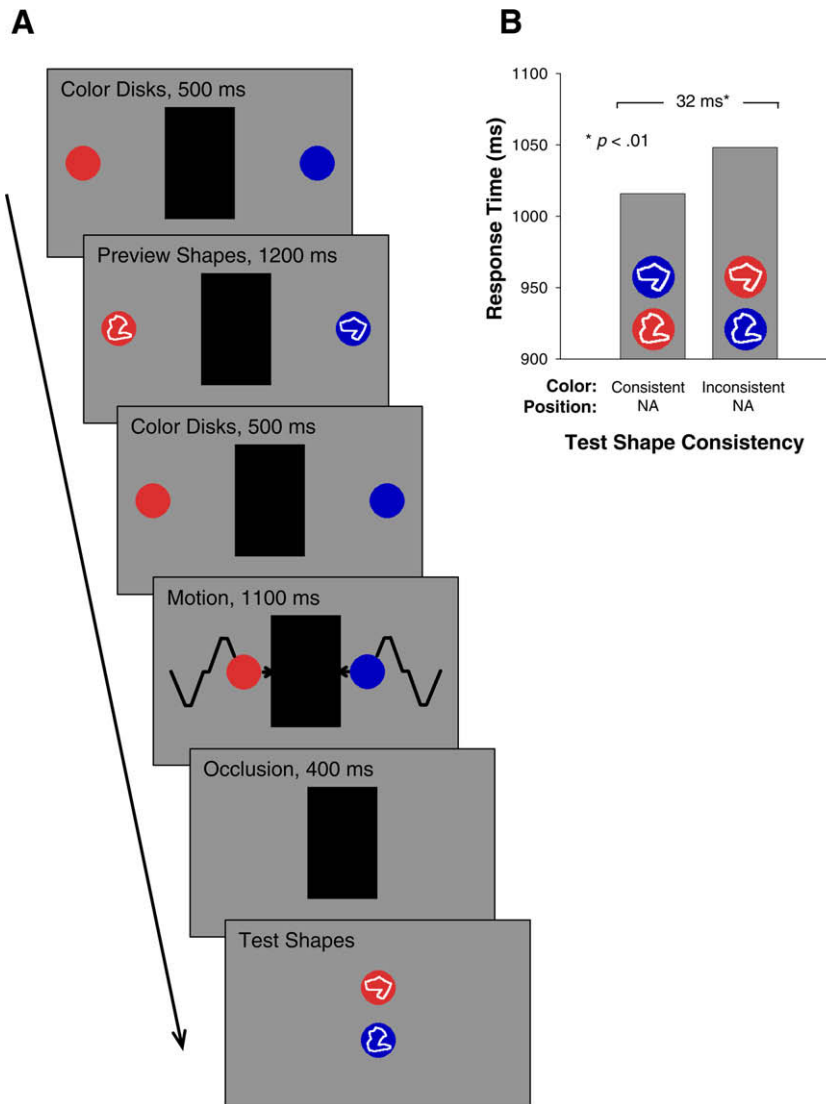
In this study, we developed a paradigm that probed the relative contributions of surface feature and spatiotemporal information to object correspondence, and we did so in a manner that eliminated any salient discontinuity on both of the two dimensions. This was achieved by introducing object manipulations during a brief period of occlusion, when they were not directly visible. In addition, we

used unfamiliar shapes as the test stimuli and required articulatory suppression throughout each trial, minimizing the possibility that participants encoded the stimuli verbally, and thus providing a more direct test of correspondence on the basis of visual object representations.

In the basic method (see Figs. 2 and 3), two colored disks were presented to the left and right of a central occluder. Novel shapes appeared in the two disks. The shapes were removed, and the empty color disks moved behind the occluder. The occluder was removed to display the two color disks containing two shapes. Both shapes were the same as in the preview, or one was replaced by a new shape. Participants reported “same” or “new”. When both test shapes were the same as in the preview, spatiotemporal consistency was manipulated (the test shapes



**Fig. 2.** (A) Sequence of events in a trial of Experiment 1. Participants reported whether the two test shapes were the same as the two preview shapes. Shape-position consistency was manipulated. Note that in all of the figures in this paper, the color disk and novel shape stimuli have been enlarged to improve visibility. In the experiments themselves, the color disks and shapes were significantly smaller relative to the size of the occluder (a total of eight disks could have fit behind the occluder without touching). (B) Mean Correct RT in each of the consistency conditions of Experiment 1. Position consistency refers to the consistency between the locations of the test shapes and the expected locations of the test shapes given the spatiotemporal history of the disks in which they appeared. The stimuli inset in the bars of the graph show the test configuration in each condition for the illustrated sample trial.



**Fig. 3.** (A) Sequence of events in a trial of Experiment 2. (In the print version of this figure, the two disk colors are represented as light grey and dark grey.) The consistency between the binding of novel shape and colored disk in the preview display and in the test display was manipulated. (B) Mean Correct RT in each of the consistency conditions of Experiment 2. The stimuli inset in the bars of the graph show the test configuration in each condition for the illustrated sample trial.

appeared in locations that were either consistent or inconsistent with the motion of the disks), and surface feature consistency was manipulated (the test shapes appeared in disks with colors that were either consistent or inconsistent with the colors of the disks in which they originally appeared). If object correspondence operations consult only spatiotemporal information, as held by the spatiotemporal dominance hypothesis, significant effects of spatiotemporal consistency should be observed, but no effects of color consistency should be observed.

The present study consisted of five experiments. In Experiment 1, we manipulated spatiotemporal consistency to ensure that our paradigm was sensitive to the spatiotemporal effects observed in previous object file experiments. In Experiments 2–5, we manipulated surface feature consistency (color), and these experiments imple-

mented increasingly stringent conditions for observing surface feature effects on object correspondence. To preview the results, we found robust effects of both spatiotemporal and color consistency on the mapping of objects across brief occlusion, demonstrating that both sources of information contribute to object correspondence. Contrary to claims of spatiotemporal dominance, surface feature consistency influenced object correspondence under a wide range of circumstances. Surface features were used to establish object correspondence when spatiotemporal information was ambiguous (Experiment 2), when spatiotemporal information was perfectly predictive of correspondence (Experiment 3), when spatiotemporal information was inconsistent with the interpretation of correspondence on the basis of a color match (Experiment 4), and even when there was salient spatiotemporal dis-

continuity directly in conflict with an interpretation of object continuity (Experiment 5).

### 3. Experiment 1

Before testing the role of surface features in object correspondence, we first sought to confirm that our modified version of the object-reviewing paradigm was sensitive to spatiotemporal consistency. The method, illustrated in Fig. 2A, was similar to the classic object-reviewing paradigm of Kahneman et al. (1992), and we expected to replicate the basic spatiotemporal consistency effect: a performance advantage when the test stimuli appeared in positions consistent with their spatiotemporal history.

Each trial began with two blue disks appearing to the left and right of a central occluder. Two novel shapes appeared in the disks briefly. Then, the empty disks moved diagonally (one up, one down), followed by horizontal motion that caused them to disappear behind the occluder at different vertical positions. The occluder was removed to reveal the two blue disks containing two test shapes (both “same” or one “new”). The test shapes appeared in positions that were either consistent or inconsistent with the expected locations given the motion of the disks. Consider the sample trial illustrated in Fig. 2. The stimuli inset in the bars of the graph show the test configuration on a “same” trial in each condition. In the consistent-position condition, the test shape that originally appeared in the left disk is presented in the bottom position, which is consistent with the movement of the left disk to the bottom position. The position of the other test shape is also consistent with the motion of the disk in which it appeared. In the inconsistent-position condition, the two test shapes appear in the inappropriate positions given the motion of their respective disks. Note that surface feature correspondence was ambiguous, because the two color disks had identical surface features.

As is standard in this literature, the data of interest came primarily from “same” trials. On “new” trials, the correspondence between the new test shape and the preview display is undefined, complicating their interpretation. Our task is equivalent to the recognition memory task of recent object file studies (Kruschke & Fragassi, 1996; Mitroff & Alvarez, 2007; Mitroff et al., 2004; Mitroff et al., 2005), except that we presented two test shapes instead of one test letter. Our method required two comparisons of test shapes with memory to successfully determine “same”, each of which could have been influenced by the surface feature or spatiotemporal properties of the object in which the test shape appeared. This method was designed to maximize consistency effects, which were more than twice as large in the present study than in previous object-file experiments.

#### 3.1. Method

##### 3.1.1. Participants

Participants in all experiments were recruited from the University of Iowa community and were between the ages of 18 and 30. Participants reported normal or corrected-

to-normal vision. They either received course credit or were paid. Twenty participants completed Experiment 1.

##### 3.1.2. Stimuli and apparatus

Stimuli are shown in Fig. 2A. The background was gray. Color disks subtended  $1.59^\circ$ . All disks were blue. Five novel shapes were created in a computer graphics program. Each shape was white and was centered within the disk. On each trial, the two shapes that appeared in the preview display were chosen randomly without replacement from the set of five. The transition between shape and disk was abrupt (i.e., not antialiased). Shapes subtended an average of  $1.16^\circ$  horizontally and  $1.14^\circ$  vertically. The central, black occluding rectangle subtended  $4.26^\circ \times 6.82^\circ$ .

The initial positions of the two color disks were centered  $7.1^\circ$  to the left and right of the screen center. The motion animation was composed of 51 frames at a rate of 20 ms/frame. The two disks first moved on a diagonal path, with one moving upward and the other downward (horizontal displacement =  $1.19^\circ$ ; vertical displacement =  $1.42^\circ$ ), assigned randomly. The diagonal motion was completed in 200 ms. Thereafter, the two objects moved on horizontal paths until they were completely occluded, which required an additional 820 ms. The vertical positions of the two objects were separated by  $2.84^\circ$  (center-to-center).

In the test image, two color disks were presented at the center of the display, separated vertically by  $2.84^\circ$ . Two shapes were presented within the disks. On ‘same’ trials, the two shapes at test were the same two shapes appearing in the preview display. On ‘new’ trials, one of the two preview shapes (randomly chosen) was replaced by a new shape (randomly chosen from the remaining three shapes). The locations of the two shapes were either consistent or inconsistent with the expected locations given the motion of the disks. Because the two disks always had the same color, color was non-informative with respect to the shape memory task.

Stimuli were displayed on a 17-in CRT monitor with a 100 Hz refresh rate. A forehead rest was used to maintain a constant viewing distance of 80 cm. Responses were collected using a serial button box. The experiment was controlled by a computer running E-Prime software.

##### 3.1.3. Procedure

On each trial, participants first saw four randomly selected digits and repeated the digits aloud throughout the trial to further minimize verbal encoding of the stimuli (the experimenter monitored digit repetition to ensure that participants complied). The participant pressed a button to begin the trial. There was a 700 ms blank delay followed by the sequence of events depicted in Fig. 2A. Because the color disks were visible before, during, and after the appearance of the novel shapes, the experience when viewing these displays was of two objects (the color disks) that temporarily acquired a new feature (the novel shape) and then moved behind the occluder. During motion and at the point of occlusion, the only means to discriminate the two objects was by virtue of their spatiotemporal properties. After the appearance of the test display, participants pressed one of two buttons to respond

that both shapes were the same as in the preview or that one shape was new. Participants were instructed to respond as quickly and as accurately as possible. In each experiment, participants were also instructed that the positions of the shapes and the colors of the disks in which they appeared were entirely irrelevant to the task, and that they should base their decision solely on the shapes themselves.

Experiment 1 was a 2 (spatiotemporal consistency)  $\times$  2 (same, new) design. Participants completed a practice block of 12 trials, drawn randomly from the full design, followed by an experiment block of 280 trials (70 in each condition). Trial order was determined randomly. The entire session lasted approximately 45 min.

### 3.2. Results and discussion

As is conventional, analyses were conducted over “same” trials, in which both test shapes had been present in the preview. “New” trials introduced a new shape, and the relationship between this test shape and the color disks was undefined. For all experiments, analyses over all trials produced the same pattern of results as analyses over same trials alone. The complete data are reported in the Appendix. RT analyses were limited to correct responses. In addition, responses more than 2.5 SD from each participant’s mean RT in each condition were eliminated as outliers (no more than 2.7% of the data in any experiment). RT trimming did not alter the pattern of results in any experiment. Mean correct RT data are reported in Fig. 2B.

The purpose of Experiment 1 was to ensure that we could observe the standard effect of spatiotemporal consistency found originally by Kahneman et al. (1992). Indeed, a robust effect of spatiotemporal consistency was obtained. Mean RT was 73 ms faster in the consistent-position condition than in the inconsistent-position condition,  $t(19) = 7.25$ ,  $p < .001$ . In addition, there was a complementary 2.2% effect on accuracy,  $t(19) = 2.25$ ,  $p = .04$ . An effect on accuracy, in addition to the relatively large effect on RT, indicates that the present method is a highly efficient means to assess object correspondence operations.

Before proceeding to examine the role of surface features in object correspondence, it is important to note that there exists some ambiguity in the object-reviewing paradigm developed by Kahneman et al. (1992) and adapted here. Correspondence is measured by the response to a test stimulus (typically a letter; in the present case, novel shapes) that is associated with the object whose continuity is being manipulated (typically a box; in the present case, colored disks). The assessment of correspondence is therefore indirect. It is possible that participants track the objects on the basis of the test feature (the letters or novel shapes) and use a test feature match to establish correspondence. The association of the test stimulus to a location could then influence the response to the test stimulus, generating the consistency effect, but without necessarily influencing the computation of object correspondence per se. The same issue applies to manipulations of surface feature consistency (Kahneman et al., 1992; Mitroff & Alvarez, 2007; present Experiments 2–5). The association of the test stimulus to a particular set of surface

features could potentially generate a consistency effect independently of the computation of correspondence.

Although possible, we think this alternative account is unlikely to explain the spatiotemporal consistency effects observed here and in earlier experiments or the surface feature consistency effects reported in Experiments 2–5. The novel shapes were clearly perceived as temporary features of otherwise stable objects. The colored disks were visible for a full second before the appearance of the preview shapes. The preview shapes disappeared 500 ms before the objects started to move. And, critically, during the motion sequence and disappearance behind the occluder (the events that created the need to establish correspondence), the shapes were not visible. It therefore does not seem likely that participants “tracked” the objects and established correspondence on the basis of a previously visible feature (the preview shape) rather than on the basis of features directly visible throughout the trial until the critical occlusion event (the spatiotemporal and surface feature properties of the objects).

## 4. Experiment 2

Experiment 2 constituted our first test of a surface feature contribution to object correspondence and thus the first test of the spatiotemporal dominance hypothesis. In the standard object-reviewing paradigm (Kahneman et al., 1992), replicated in Experiment 1, spatiotemporal consistency is manipulated in the context of ambiguous surface feature information, because the two objects have identical surface features. Spatiotemporal information is therefore the only available cue to correspondence. In Experiment 2, we developed an analogous test of surface feature contributions to object correspondence: Surface feature (color) consistency was manipulated in the presence of ambiguous spatiotemporal information, and surface feature information was the only available cue to correspondence. The method is illustrated in Fig. 3A.

In the preview display, the two disks differed in color. After the presentation of the preview shapes, the two color disks moved initially on quasi-random paths. Then, both objects disappeared behind the occluder while traveling horizontally at the *same* vertical position. In the test display, the two disks were separated vertically, as in Experiment 1. The presentation of vertically separated items at test was consistent with the history of the display, because the objects had a history of moving up or down erratically. However, spatiotemporal correspondence was ambiguous, because the spatiotemporal history of the display did not specify the relative locations of the two objects in the test display. Of central importance, the colors of the disks in which the test shapes appeared were either consistent or inconsistent with the color-shape pairings in the preview. Consider the sample trial in Fig. 3. In the color-consistent condition, the shape that originally appeared within the red disk (depicted in light gray in the print version of the figure) is presented within the red disk at test. Color is also consistent for the shape that originally appeared in the purple disk (depicted in dark grey in the print version of the figure). In the color-inconsistent condition, however,



the pairings between test shapes and color discs are reversed, and both test shapes appear within colors that are inconsistent with their preview colors.

If surface features can serve to establish object correspondence across brief occlusion, a performance advantage should be observed for the color-consistent condition over color-inconsistent condition. Such evidence would demonstrate that correspondence operations are not limited to spatiotemporal features and that surface features are consulted, at least under conditions of spatiotemporal ambiguity.

#### 4.1. Participants

Twenty new participants completed Experiment 2.

#### 4.2. Stimuli, apparatus, and procedure

The method in Experiment 2 was the same as that in Experiment 1, with the following exceptions. In Experiment 2, the colors of the two disks were different. Each color was chosen randomly without replacement from a set of five colors: blue, green, red, purple, and yellow. In the test display, the colors of the disks in which the test shapes appeared were either consistent or inconsistent with the preview colors. In the latter case, the pairings of novel shape and color were swapped from preview to test.

The motion sequence consisted of a series of diagonal and horizontal motion components. Each diagonal component displaced the object  $0.91^\circ$  horizontally and  $2.3^\circ$  vertically over 200 ms of animation. To smooth the transition between diagonal components, each diagonal component was separated by  $0.23^\circ$  of horizontal motion over 40 ms. The motion sequence started with the first diagonal component, with the vertical direction (up or down) randomly chosen for each object. The second diagonal component brought the object back to the vertical midline. The direction of the third diagonal component (up or down) was again randomly chosen for each object. And the fourth diagonal component again brought the objects back to the vertical midline. The last component was horizontal motion of  $1.59^\circ$  over 180 ms, which ended with the full occlusion of both objects at the vertical midline. This entire motion sequence required 1100 ms. What the observer perceived was the unpredictable upward and downward diagonal motion of the objects, followed by horizontal motion that caused the objects to be occluded at the same vertical position. After the 1100 ms motion sequence, there was a 400 ms period of occlusion before the test display. This period was easily sufficient to support the possibility of another diagonal motion behind the occluder, consistent with the appearance of the test objects at different vertical positions.

The test positions of the two color disks were randomly chosen on each trial. The pairings of test shapes and color disks were either consistent with the shape-color pairings in the preview display or inconsistent with those pairings.

#### 4.3. Results and discussion

Mean correct RT data are reported in Fig. 3B. Mean RT was 32 ms faster in the color-consistent condition than in the color-inconsistent condition,  $t(19) = 3.25$ ,  $p = .004$ .

There was a complementary, but nonsignificant, 0.9% effect on accuracy,  $t(19) = 0.83$ ,  $p = .42$ .

The results of Experiment 2 demonstrate that surface feature consistency can be used to establish object correspondence when spatiotemporal correspondence is ambiguous. They stand in contrast with the results of Mitroff and Alvarez (2007) and Kahneman et al. (1992), in which no effect of surface feature consistency was observed. The potential causes of the difference between our results and those of previous studies are examined further in Experiment 5.

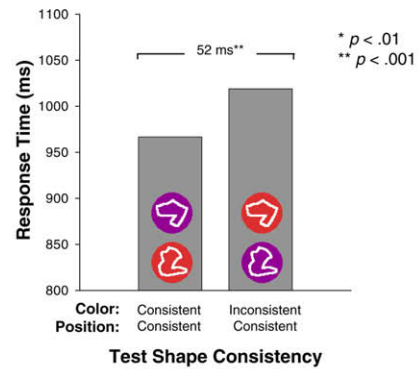
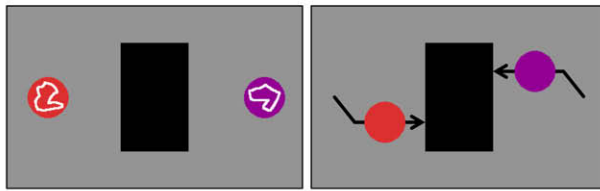
The surface feature consistency effect observed in Experiment 2 eliminates a strong version of the spatiotemporal dominance hypothesis holding that only spatiotemporal information is consulted by correspondence operations (Kahneman et al., 1992). However, the results of Experiment 2 are potentially consistent with a weaker version of the spatiotemporal dominance hypothesis. In this view (Flombaum et al., 2009; Scholl, 2007), surface features can be consulted in the computation of object correspondence, but only when spatiotemporal information is ambiguous or non-informative. That is, surface features may be used only as a “last resort” if correspondence cannot be established on the basis of spatiotemporal cues. When spatiotemporal and surface feature information are both potentially informative, spatiotemporal information will “trump” surface feature information in correspondence operations, and no effect of surface feature consistency should be observed. In Experiments 3 and 4, we tested this weaker version of the spatiotemporal dominance hypothesis by manipulating surface feature consistency in the presence of predictive, rather than ambiguous, spatiotemporal information.

### 5. Experiment 3

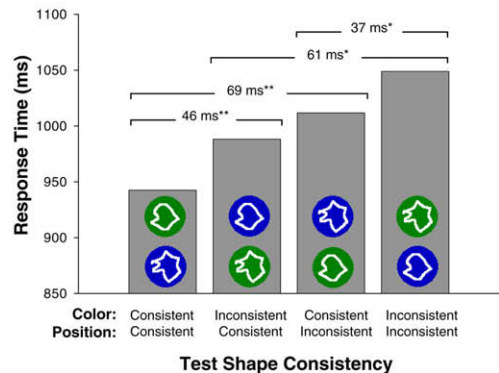
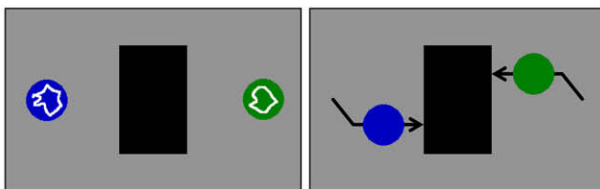
In Experiment 3, the color manipulation of Experiment 2 was combined with the unambiguous motion used in Experiment 1 (see Fig. 4). Two differently colored disks appeared to the left and right of the occluder. They then moved behind the occluder at different vertical positions (as in Experiment 1). Spatiotemporal history served to distinguish the two objects and could have been used to establish correspondence after occlusion. In addition, the test shapes always appeared in locations consistent with the motion history of the two disks, so spatiotemporal information also perfectly predicted the locations of the test shapes. Color consistency was manipulated within this predictive spatiotemporal context. In the color-consistent condition, the test shapes appeared in disks that matched the colors of the preview disks in which the shapes originally appeared. In the color-inconsistent condition, both test shapes appear within colors that were inconsistent with their preview colors.

In this method, correspondence could have been established on the basis of spatiotemporal information alone. Thus, the hypothesis that surface features are consulted only when spatiotemporal information is ambiguous (Flombaum et al., 2009; Scholl, 2007) predicts that spatiotemporal information should dominate correspondence

### A. Experiment 3: Color Consistency



### B. Experiment 4: Spatiotemporal Consistency X Color Consistency



**Fig. 4.** Experimental design and mean correct RT data for Experiments 3 (A) and 4 (B). (In the print version of this figure, the two disk colors are represented as light grey and dark grey.) For each experiment, the left-hand portion of the figure shows the key events in a sample trial. Color consistency refers to the consistency between the binding of novel shape and colored disk in the preview display and in the test display. Position consistency refers to the consistency between the locations of the test shapes and the expected locations of the test shapes given the spatiotemporal history of the disks in which they appeared. The stimuli inset in the bars of the graph show the test configuration in each condition for the illustrated sample trial.

operations, and color should not be consulted. If so, then the consistent-color advantage observed in Experiment 2 should be eliminated in Experiment 3.

#### 5.1. Method

##### 5.1.1. Participants

Twenty new participants completed Experiment 3.

##### 5.1.2. Stimuli, apparatus, and procedure

The motion sequence was identical to that in Experiment 1. In the test display, the test shapes always appeared in the locations consistent with the motion of the disks. As in Experiment 2, the colors of the two disks differed, and color consistency was manipulated at test. In the test display, the colors of the disks in which the test shapes appeared were either consistent or inconsistent with the preview colors. In the latter case, the pairings of novel shape and color were swapped from preview to test.

#### 5.2. Results and discussion

Mean correct RT data are reported in Fig. 4A. Contrary to the claim that object correspondence operations are dominated by spatiotemporal information when spatiotemporal information is unambiguous, a robust color consistency effect was observed. Mean RT was 52 ms faster

in the color-consistent condition than in the color-inconsistent condition,  $t(19) = 4.42$ ,  $p < .001$ . There was a complementary 3.5% effect on accuracy,  $t(19) = 2.94$ ,  $p = .008$ .

To ensure that color consistency effects were not caused by differences in the contrast between the shape and the color on which the shape appeared (a possibility, because colors differed in luminance), we conducted a control experiment ( $N = 20$ ). Color disks were made larger ( $2.57^\circ$  diameter), and the shapes always appeared within a dark gray disk ( $1.59^\circ$ ) centered within the larger color disk, controlling shape-background contrast (see Fig. 5). Otherwise, the method was identical to that in Experiment 3. A robust color consistency effect was again observed. Mean RT was 42 ms faster in the color-consistent condition (885 ms) than in the color-inconsistent condition (927 ms),  $t(19) = 2.82$ ,  $p = .01$ , with a complementary 3.6% effect on accuracy,  $t(19) = 4.26$ ,  $p < .001$ .

In Experiment 3, spatiotemporal history distinguished the locations of the objects after occlusion and thus was informative of correspondence. In addition, the test shapes always appeared in the appropriate locations given the motion of the disks. Spatiotemporal information perfectly predicted test shape, whereas color did not predict test shape (because on half the trials, color-shape binding was consistent, and on the other half, it was inconsistent). If spatiotemporal properties “trump” surface feature properties when spatiotemporal properties are informa-

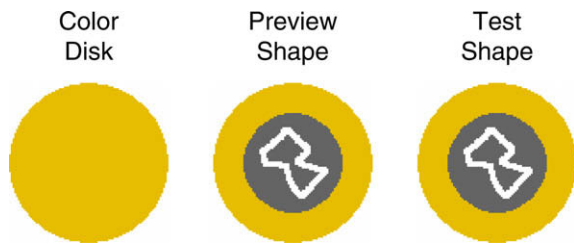


Fig. 5. Sample stimuli used in the Experiment 3 control.

tive, correspondence should have been established on the basis of spatiotemporal information in Experiment 3, and color should not have influenced correspondence. Yet, a robust effect of color consistency was observed. This result cannot be accommodated by a theory holding that, in general, spatiotemporal information trumps surface feature information in correspondence operations (although it is still possible that spatiotemporal information might trump surface feature information in other forms of correspondence, a topic addressed in the General Discussion).

## 6. Experiment 4

In Experiment 4, we expanded the design of Experiment 3 to manipulate both spatiotemporal and color consistency independently (see Fig. 4B), allowing us to assess the relative contributions of surface feature and spatiotemporal information to object correspondence in this paradigm. In the preview display, the novel shapes appeared within two disks that differed in color. In the test display, color and position were either both consistent, one consistent and the other inconsistent, or both inconsistent.

The design of Experiment 4 provides two means to assess the contributions of spatiotemporal and surface feature information to object correspondence. Similarly to Experiments 1–3, each of the conditions in which only one of the two dimensions is inconsistent (two middle bars of Fig. 4B) can be compared against the condition in which both are consistent (left-most bar of Fig. 4B). These contrasts assess the performance *decrement* associated with introducing inconsistency on one of the two dimensions. We expected to replicate the results of Experiments 1–3. The second method is to compare each of the conditions in which only one of the two dimensions is consistent (two middle bars of Fig. 4B) against the condition in which both are *inconsistent* (right-most bar of Fig. 4B). These contrasts assess the performance *benefit* associated with introducing consistency on one of the two dimensions.

The latter type of comparison (against the both-inconsistent baseline) allowed us to test the effect of consistency on one dimension when the other dimension was inconsistent with the interpretation of correspondence. This provided a particularly strong test of the role surface features in object correspondence. If an effect of color consistency were to be observed in this comparison, it would indicate a contribution of surface feature information to correspondence operations even when spatiotemporal information was inconsistent with the interpretation of correspondence. Likewise, an effect of spatiotemporal consistency

would indicate a contribution of spatiotemporal information to correspondence operations even when surface feature information was inconsistent with correspondence.

### 6.1. Method

#### 6.1.1. Participants

Forty new participants completed Experiment 4.

#### 6.1.2. Stimuli, apparatus, and procedure

The method was identical to Experiment 3, except that the test shapes could appear either at the consistent or inconsistent positions given the spatiotemporal history of the objects, and the color of the disks in which the test shapes appeared was either consistent or inconsistent with the shape-color pairing in the preview display. Experiment 3 was a 2 (spatiotemporal consistency)  $\times$  2 (color consistency)  $\times$  2 (same, new) design. Participants completed a practice block of eight trials followed by an experiment block of 336 trials (42 in each condition). The entire session lasted approximately 55 min.

### 6.2. Results and discussion

Mean correct RT data are reported in Fig. 4B.

We first compared each of the conditions in which one of the two dimensions was inconsistent against the both-consistent condition. Replicating Experiment 1, mean RT was 69 ms faster in the both-consistent condition than in the position-inconsistent/color-consistent condition,  $t(39) = 3.75$ ,  $p < .001$ , with a complementary, but nonsignificant, 1.4% effect on accuracy,  $t(39) = 1.29$ ,  $p = .20$ . Replicating Experiments 2 and 3, mean RT was 46 ms faster in the both-consistent condition than in the position-consistent/color-inconsistent condition,  $t(39) = 4.47$ ,  $p < .001$ , with a complementary, but nonsignificant, 1.3% effect on accuracy,  $t(39) = 1.23$ ,  $p = .22$ . The magnitude of the position effect (69 ms) and color effect (46 ms) did not differ,  $t(39) = 1.38$ ,  $p = .18$ .

Next, we examined whether consistency on one dimension would influence correspondence despite evidence inconsistent with correspondence on the other dimension. Each of the conditions in which one of the two dimensions was consistent was compared against the both-inconsistent condition. There was a reliable effect of spatiotemporal consistency. Mean RT was 61 ms faster in the position-consistent/color-inconsistent condition than in the both-inconsistent condition,  $t(39) = 3.47$ ,  $p = .001$ , with a complementary 3.0% effect on accuracy,  $t(39) = 2.78$ ,  $p = .008$ . In addition, there was a reliable effect of color consistency. Mean RT was 37 ms faster in the position-inconsistent/color-consistent condition than in the both-inconsistent condition,  $t(39) = 2.88$ ,  $p = .006$ , with a complementary 2.9% effect on accuracy,  $t(39) = 2.55$ ,  $p = .01$ .

The results of Experiment 4 demonstrate that the effect of consistency on one dimension is not strongly constrained by consistency on the other dimension. An effect of color consistency was observed when spatiotemporal information was consistent with correspondence, and an effect of color consistency was observed when spatiotemporal information was inconsistent with correspondence.

Similarly, an effect of spatiotemporal consistency was observed regardless of color consistency. An alternative way to frame this point is to note that the combined effect of color and spatiotemporal consistency was larger than the effect of either type of consistency alone.

In addition to this main finding, Experiment 4 eliminates a potential concern with the design of Experiment 3. In Experiment 3, performance may have been influenced by the consistency between the pre-occlusion and post-occlusion positions of the colors themselves. In the color-consistent condition of Experiment 3, the colors appeared in the appropriate locations given the spatiotemporal history of the objects. However, in the color-inconsistent condition, the colors appeared in locations that were inconsistent with the spatiotemporal history of the objects. Consider Fig. 4A. In the color-consistent condition, the red and purple disks (light grey and dark grey in the print version of the figure) are in the appropriate locations given the motion of the objects, but in the color-inconsistent condition, they are not. (Of course, the test shapes themselves were always in the appropriate locations.) It is therefore possible that the effect in Experiment 3 was driven not by color-shape consistency but rather by color-position consistency. Note that this alternative account would still be problematic for the spatiotemporal dominance hypothesis, because, under that view, color should not determine correspondence at all.

Nevertheless, the results of Experiment 4 eliminate this alternative explanation in terms of color-position consistency. Two conditions of Experiment 4 are relevant, the position-inconsistent/color-consistent condition and the both-inconsistent condition (two right-hand bars of Fig. 4B). These two conditions differ in the consistency of the shape-color pairings; shape-position pairing is inconsistent in both. Comparison of these two conditions therefore probes the effect of color consistency against a position-inconsistent baseline. As reported above, there was a reliable 37 ms effect of color-shape consistency on RT, and a reliable 2.9% effect on accuracy. This effect cannot be explained by differences in the position of the colors themselves. In the both-inconsistent condition, the colors appeared in the *appropriate* locations given motion before occlusion. In the position-inconsistent/color-consistent condition, the colors appeared in the *inappropriate* locations. Yet, a reliable advantage was found for the latter condition over the former, eliminating the possibility that it was the positions of the colors that was driving the effect. We can therefore be confident that the color effects in Experiments 3 and 4 were caused by the consistency between color and test shape.

Finally, we have interpreted the results of Experiment 4 as indicating that both spatiotemporal and surface feature consistency contributed to the computation of object correspondence. It is possible, however, that correspondence was established initially on the basis of the spatiotemporal features of the objects, and that the effects of color consistency were generated later, only after spatiotemporal correspondence had already been computed. Similarly, it is possible that correspondence was established initially on the basis of surface feature cues, and that the effects of spatiotemporal consistency were generated only after corre-

spondence had already been established. Although these alternatives are possible within the context of Experiments 3 and 4 (in which both types of cues were potentially informative of correspondence), the effects of spatiotemporal and color consistency in Experiment 1 and 2 could not have been generated after correspondence had already been computed, because in these experiments, spatiotemporal cues (Experiment 1) and color cues (Experiment 2) were the only cues available upon which to establish correspondence in the first place. Thus, the full data set indicates that both spatiotemporal and surface feature cues can contribute to the computation of object correspondence and that effects of consistency on the two dimensions are largely independent.

## 7. Experiment 5

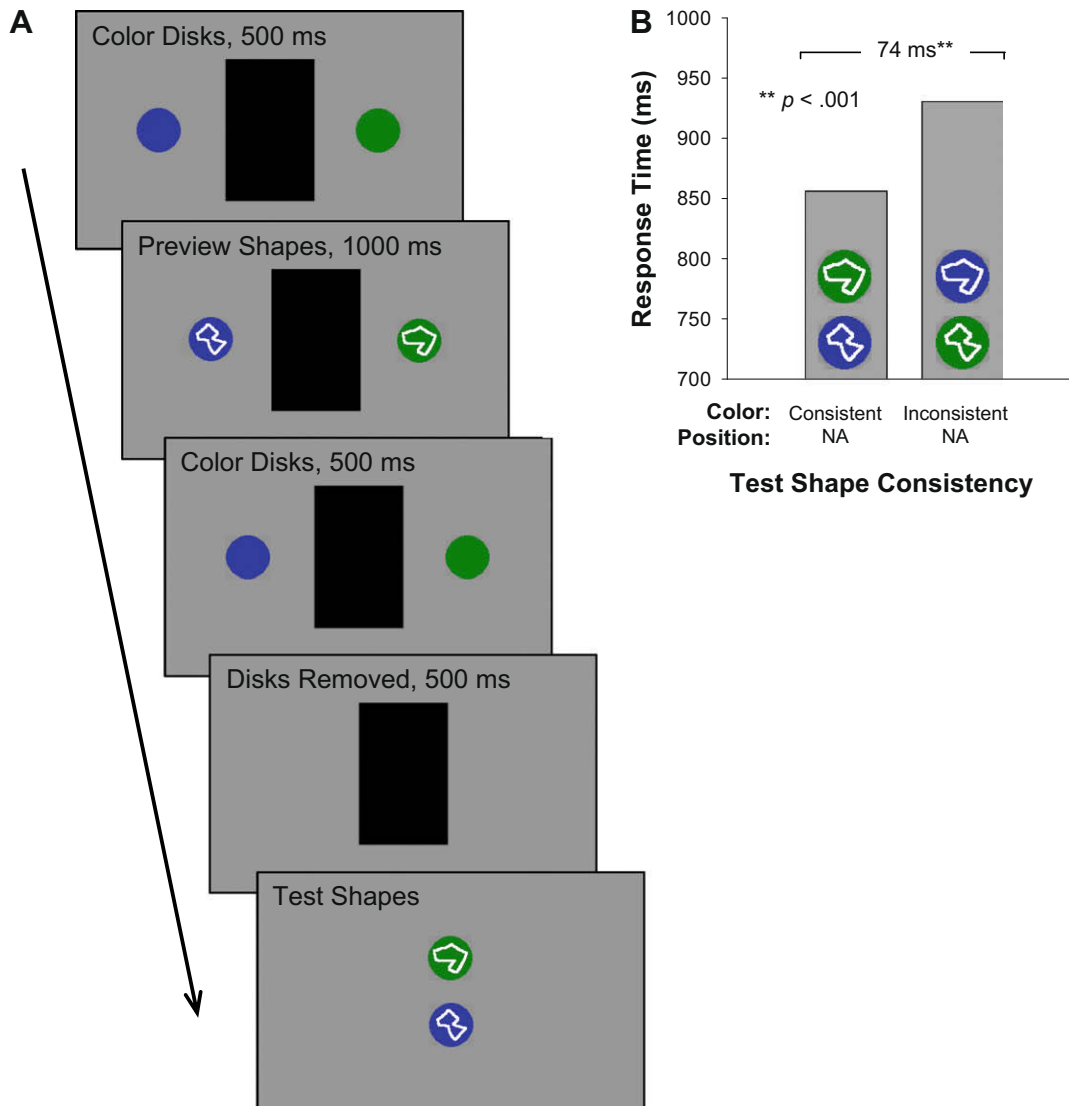
In the experiments so far, the objects always moved behind the occluder, and thus there were clear perceptual cues to indicate that the objects should be found behind the occluder in the test display. It is possible that the effects of color consistency were observed in the present experiments because motion provided a plausible spatiotemporal account of how the objects came to be found behind the occluder. That is, surface features may contribute to object correspondence only when there is coherent spatiotemporal information to account for their general locations.<sup>2</sup>

To address this possibility, in Experiment 5 we eliminated the motion information from the paradigm, eliminating any plausible spatiotemporal account of the difference in location between the preview and test objects. Experiment 5 therefore constituted a particularly stringent test of surface feature correspondence: Will surface feature consistency be sufficient to override spatiotemporal information that is directly at odds with an interpretation of object continuity?

The Experiment 5 method (Fig. 6A) was essentially the same as in previous experiments, except there was no object motion. The color disks disappeared from their positions flanking the occluder, and 500 ms later, the occluder was removed to reveal the color disks at the center of the screen. There was no plausible spatiotemporal account for the change in location, because there was no motion to link the preview and test positions of the color disks. Color consistency was again manipulated. If surface features are consulted only when spatiotemporal information provides a plausible link between the original and new locations of the objects, no effect of color consistency should be observed.

The elimination of motion generated a paradigm very similar to the paradigm of Mitroff and Alvarez (2007); see Fig. 1. Just as in Mitroff and Alvarez, the two objects were moved to new locations without any linking motion. That is, the Experiment 5 paradigm introduced a salient spatiotemporal discontinuity, allowing us to test whether the presence/absence of spatiotemporal discontinuity is the primary cause of the difference between our findings and those of Mitroff and Alvarez.

<sup>2</sup> We thank an anonymous reviewer for suggesting this possibility.



**Fig. 6.** (A) Sequence of events in a trial of Experiment 5. (In the print version of this figure, the two disk colors are represented as light grey and dark grey.) (B) Mean correct RT in each of the consistency conditions of Experiment 5. The stimuli inset in the bars of the graph show the test configuration in each condition for the illustrated sample trial.

### 7.1. Participants

Twenty new participants completed Experiment 5.

### 7.2. Stimuli, apparatus, and procedure

The method was the same as that in Experiment 2, with the following exceptions. The initial positions of the two color disks were centered  $4.3^\circ$  to the left and right of the screen center. After the preview shapes had been removed, the color disks remained visible for 500 ms. They were then removed for 500 ms (occluder only), followed by the test display.

### 7.3. Results and discussion

Mean correct RT data are reported in Fig. 6B. Mean RT was 74 ms faster in the color-consistent condition than in the col-

or-inconsistent condition,  $t(19) = 5.99$ ,  $p < .001$ . There was a complementary 5.3% effect on accuracy,  $t(19) = 3.30$ ,  $p = .003$ .

In Experiment 5, a highly robust color consistency effect was observed despite the fact that there was no spatiotemporal link between the preview and test locations of the objects, and thus no plausible spatiotemporal account of how the objects came to be behind the occluder. It appears that surface features can establish object correspondence even when there is direct spatiotemporal information that is inconsistent with object continuity. This in turn suggests that the critical difference between the present study and that of Mitroff and Alvarez is not necessarily the presence of a salient spatiotemporal discontinuity, because Experiment 5 revealed a color consistency effect despite salient spatiotemporal discontinuity.

It is difficult to determine precisely why we found surface feature consistency effects, whereas Mitroff and Alva-

red did not. We think the most plausible explanation is that the present paradigm was simply more sensitive to consistency manipulations (both of spatiotemporal and surface feature information) than was the method of Mitroff and Alvarez. First, novel shapes and articulatory suppression in the present study minimized verbal encoding of the test stimuli and maximized the possibility that the features of the objects and the test stimuli were integrated into a coherent object representation in VWM (Luck & Vogel, 1997), increasing sensitivity to manipulations of the binding between object features and test shapes. Second, the presentation of two test stimuli, instead of one, potentially multiplied the effects of consistency compared with previous experiments using only a single test stimulus. This increased sensitivity was observed not only in the ability to detect effects of surface feature consistency but also in the magnitude of the spatiotemporal consistency effects, which were approximately three times as large in the present study as in Mitroff and Alvarez (and other experiments using letters and a single test stimulus).

## 8. General discussion

In the present experiments, we examined the contributions of spatiotemporal and surface feature information to the computation of object correspondence across a brief period of occlusion. The object-file framework of Kahneman et al. (1992) holds that object representations are addressed by their spatiotemporal properties and that spatiotemporal information is primary in determining object correspondence, a *spatiotemporal dominance hypothesis*. In five experiments modeled after the Kahneman et al. object-reviewing paradigm, we observed that object correspondence operations were not necessarily restricted to spatiotemporal cues. A salient surface feature cue, color, was used to establish object correspondence across brief occlusion. Robust effects of color consistency on object correspondence were observed across a series of experiments that instituted progressively more stringent demands on surface feature correspondence. In Experiment 2, object correspondence was established on the basis of surface feature cues when spatiotemporal information was ambiguous and color was the only feature available to distinguish the objects. In Experiment 3, color was consulted in object correspondence operations despite unambiguous spatiotemporal information that could have been used to compute correspondence. In Experiment 4, the contributions of spatiotemporal and surface feature information to object correspondence were found to be largely independent. In addition, surface features were consulted even on trials when spatiotemporal information was inconsistent with the interpretation of correspondence on the basis of color. Finally, in Experiment 5, effects of surface feature consistency were observed despite the absence of motion to link the pre- and post-occlusion objects. Thus, color can be used to establish object correspondence even in the presence of spatiotemporal information that directly contradicts an interpretation of object continuity.

Together, these results provide clear evidence that correspondence operations across brief occlusion consult surface feature information and are not limited to spatiotemporal information. The results cannot be accommodated by theo-

ries claiming that, as a general rule, correspondence operations consult only spatiotemporal information (Kahneman et al., 1992), and they cannot be accommodated by theories claiming that, as a general rule, spatiotemporal features “trump” surface features when the former are informative of objecthood (Flombaum et al., 2009; Scholl, 2007).

### 8.1. Converging evidence from other perceptual domains

The conclusions from the present study apply to the domain of brief occlusion events. Converging evidence of surface feature correspondence has been found recently in other domains. Richard, Luck, and Hollingworth (2008; see also Hollingworth et al. (2008)) examined correspondence operations across saccadic eye movements. Participants generated a saccade to one of two objects. During the saccade, the objects were shifted so that the eyes landed between them, which simulated the common circumstance that the eyes fail to land on a target object. Participants were instructed to correct their gaze to the object that was either in the appropriate location (e.g., “top”) or had the appropriate surface feature property (e.g., “red”). On a subset of trials, the two objects switched properties so that the correction had to be made to (1) an object that was in the correct location but had the wrong color or (2) to an object that had the correct color but was in the wrong location. In accordance with the results of the present study, both types of inconsistency impaired participants’ ability to correct gaze to the appropriate object, suggesting that both spatial and surface feature properties are used to map objects across saccades. The Richard et al. data show that surface features are functional in object correspondence for the most frequent and the briefest form of disruption introduced by dynamic vision (saccades occur, on average, three times per second and last only 20–60 ms). In addition, Richard et al. found that the weighting of surface feature and position information was governed by the demands of saccade target selection. If the saccade target was selected on the basis of its position, position cues were weighted more heavily than surface feature cues in correspondence operations after the saccade. However, if the saccade target was selected on the basis of its surface features, surface feature cues were weighted more heavily than spatial cues in correspondence operations after the saccade.

It is still possible that there are other domains in which spatiotemporal information is dominant. As reviewed above, the most likely candidate domain is motion correspondence (Flombaum et al., 2009). In apparent motion experiments using displays composed of multiple objects, spatiotemporal factors, such as inter-object distance, tend to drive object mapping. Surface feature information can influence the perception of apparent motion, but this typically occurs only when spatial properties are non-informative (Dawson, 1991). In addition, motion is perceived between two objects even when those objects differ on salient surface feature attributes (Kolers, 1972). Yet, recent evidence suggests a role for surface features in the computation of apparent motion, even in the presence of unambiguous spatial information. Moore et al. (2007) presented a dynamic object display in which an object disk was shifted in a series of steps around a virtual circle. Par-

ticipants typically perceived a single object moving in a circular path. In one of the frames, the size of the object was changed. The position of the changed object was still consistent with the interpretation of a single object moving on a circular path. However, instead of perceiving a single object, participants perceived two objects simultaneously: a trace of the original object at the previous location and a second (larger) object at the current location. Similar effects have been found for color change (Moore & Enns, 2004). Thus, when color and size were inconsistent with the interpretation of continuity, the perception of a single, continuous object was blocked. Although this finding indicates that surface feature differences can control apparent motion under some circumstances, it is still the case that, in general, spatiotemporal factors appear to be weighted more heavily than surface features in apparent motion.

Another possibility is that spatiotemporal information is dominant in the conscious perception of object persistence (Scholl, 2007). In this view, operations that lead to the conscious experience of a single, persisting object consult spatiotemporal information preferentially. The present data cannot speak to this issue, because in the object-reviewing paradigm, correspondence is assessed indirectly (by means of association with a test stimulus) and conscious experience of object continuity is not probed (Mitroff et al., 2005). However, the Moore et al. (2007) data provide a demonstration proof that conscious perception of object persistence can be controlled by surface feature consistency. In addition, humans consciously perceive object persistence across eye movements (an object viewed across a saccade is perceived as continuously present), and correspondence operations across saccades can be controlled by surface feature attributes (Hollingworth et al., 2008; Richard et al., 2008). Thus, we can safely conclude that although there may be some circumstances under which spatiotemporal information is weighted more heavily than surface feature information, there does not appear to be any perceptual domain in which spatiotemporal information is the sole input to object correspondence operations.

## 8.2. Object representations as addressable both by spatiotemporal features and by surface features

The original claim of spatiotemporal dominance was informed by a consideration of the computational benefits of addressing objects by their locations (Kahneman et al., 1992). These benefits are considerable. There may be multiple cars in view, but there will typically only be one object occupying any given location. Addressing objects by their locations therefore allows the visual system to specify a particular individual object (a token) and distinguish it from other objects with similar visual or conceptual features (Kahneman et al., 1992). In addition, by addressing an object by its unique spatial location, the visual system can ensure that features of the object at that location, and only the features of that object, are bound together through shared position (Treisman, 1988). In this manner, position can establish “objecthood”, enabling a coherent, object-based representation to be formed. Further, addressing objects by their locations could simplify object

reference in a variety of cognitive operations (Ballard, Hayhoe, Pook, & Rao, 1997; Pylyshyn, 2000). Instead of referring to a complex set of surface feature properties when implementing an action, such as “follow [white, 4-door, boxy car]”, one need only refer to the index marking that object, “follow [index]”. As long as the spatial index can be maintained on the appropriate object, that object can be followed without any reference at all to its surface feature properties (as is likely to be the case in multiple-object tracking studies, Pylyshyn, 2004; Pylyshyn & Storm, 1988). Finally, addressing objects by their locations could alleviate limitations on memory for the surface feature properties of objects. The phenomenon of change blindness has led some researchers to conclude that very little perceptual detail is remembered from objects. Marking objects by their locations would allow one to efficiently direct attention back to the appropriate object location when information about that object is required by the task (Pylyshyn, 2000; Spivey, Richardson, & Fitneva, 2004).

There is validity to all of these arguments. And it is likely that there do exist forms of object representation that are addressed by spatial location, particularly in dorsal stream systems that have been implicated in inherently spatial tasks such as multiple-object tracking (Culham et al., 1998) and that have limited access to the surface feature properties of objects. Our argument is that object representations also can be addressed by their surface feature properties, which is likely to occur in ventral stream systems with precise coding of object surface features but with limited access to the spatiotemporal properties of objects. In addition, we argue that the visual system makes use of such object representations to individuate object tokens and to establish object correspondence across brief perceptual disruption.

In order to do so, surface feature representations of objects would need to retain sufficient object detail to distinguish individual tokens of a particular object type. Contrary to claims that surface feature memory is impoverished (Pylyshyn, 2000; Spivey et al., 2004), token-specific details of complex objects are reliably preserved in visual working memory (VWM) (Gordon et al., 2008; Henderson & Hollingworth, 2003; Henderson & Siefert, 2001; Tarr, Bulthoff, Zabinski, & Blanz, 1997), even after attention has been directed elsewhere (Gajewski & Brockmole, 2006; Hollingworth, 2004; Johnson, Hollingworth, & Luck, 2008). And VWM representations can retain quite precise surface feature information, supporting subtle, within-category color discriminations (Zhang & Luck, 2008). Thus, brief forms of visual memory are certainly capable of retaining surface feature information of sufficient precision to individuate object tokens and to establish object correspondence.

Because VWM can maintain token-specific surface feature information, a spatial index is not necessarily required to individuate tokens of a particular conceptual type, undermining one of the key motivations of the object-file claim of spatiotemporal dominance (Kahneman et al., 1992). For example, a representation of the surface feature properties of the white, 4-door, boxy car would be perfectly sufficient to distinguish that token of a car from most other cars. One would only require spatial indexes to individuate object tokens when there were multiple,

highly similar objects or a set of identical objects visible. Although laboratory tasks probing dynamic vision often create precisely this circumstance (Kahneman et al., 1992; Pylyshyn & Storm, 1988; Treisman & Gelade, 1980), a scene composed of multiple, identical objects is certainly uncommon in the real world. Under normal circumstances, a representation of the surface features of an object will be sufficient to distinguish it from other visible objects.

We think it is unlikely that the visual system would ignore such highly informative surface feature differences between objects in favor of establishing correspondence on the basis of spatiotemporal information alone. If, after being distracted by the pedestrian, a red, 2-door, convertible were to be present in the expected location of the white sedan, we doubt that many people would make an error in object correspondence and follow the red car to the wrong destination. Instead, surface feature memory would contribute to the correspondence operation, and the white car would be successfully followed, despite some degree of perceived inconsistency in spatial information. The present data demonstrate that, indeed, memory for surface features is used to establish object correspondence across brief perceptual disruption, even in the presence of spatiotemporal information that is inconsistent with the interpretation of correspondence. In general, we propose that since spatio-

temporal and surface features can both contribute to the mapping of objects across brief disruption, both will be utilized by the visual system, and the contribution of the two sources of information will likely vary as a function of utility within the current task (see also Richard et al., 2008). If there are multiple similar objects visible, then spatial position could be the primary means to individuate objects and map them across disruptions. However, if objects have clear surface feature differences, surface feature information can provide an efficient means to individuate them and map them across disruption.

Neither spatiotemporal nor surface feature information will always provide a reliable solution to the correspondence problem, however. Both are potentially prone to error. For example, if an object changes motion direction during a brief perceptual disruption, correspondence on the basis of spatiotemporal properties might certainly fail. Similarly, if an object were to rotate during a perceptual disruption, revealing a different set of surface features, correspondence on the basis of surface features might fail. The use of both sources of information, as observed in the present study, maximizes the likelihood that the visual system will generate a successful solution to the problem of object correspondence.

	Position consistent		Position inconsistent					
	Same	New	Same	New				
<i>Experiment 1</i>								
Correct RT (ms)	855	886	928	907				
Accuracy (%)	91.0	83.9	88.8	85.9				
	Color consistent		Color inconsistent					
	Same	New	Same	New				
<i>Experiment 2</i>								
Correct RT (ms)	1016	1044	1048	1039				
Accuracy (%)	90.3	80.0	89.4	81.4				
<i>Experiment 3</i>								
Correct RT (ms)	967	969	1019	978				
Accuracy (%)	93.9	88.1	90.4	90.7				
<i>Experiment 3 control</i>								
Correct RT (ms)	885	957	927	953				
Accuracy (%)	89.6	76.8	86.0	78.9				
	Both consistent		Color inconsistent		Position inconsistent		Both inconsistent	
	Same	New	Same	New	Same	New	Same	New
<i>Experiment 4</i>								
Correct RT (ms)	942	1007	988	1040	1012	1030	1049	1019
Accuracy (%)	93.1	82.9	91.8	84.2	91.7	81.9	88.8	83.8
	Color consistent		Color inconsistent					
	Same	New	Same	New				
<i>Experiment 5</i>								
Correct RT (ms)	857	940	931	962				
Accuracy (%)	92.4	73.3	87.1	75.4				



## Acknowledgements

The research was supported by NIH Grant R01EY017356. We would like to thank Cathleen Moore, Steve Luck, and Shaun Vecera for helpful discussions of the research.

## Appendix

Mean percentage correct and RT data from Experiments 1–5 and the Experiment 3 control.

## References

- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–767.
- Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. H. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology*, *80*, 2657–2670.
- Currie, C. B., McConkie, G. W., Carlson-Radvansky, L. A., & Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. *Perception and Psychophysics*, *62*, 673–683.
- Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review*, *98*, 569–603.
- Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*, 295–313.
- Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 135–164). Oxford University Press.
- Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin and Review*, *13*, 581–587.
- Gordon, R. D., Vollmer, S. D., & Frankl, M. L. (2008). Object continuity and the transsaccadic representation of form. *Perception and Psychophysics*, *70*, 667–679.
- Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, *123*, 410–426.
- Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics*, *65*, 58–71.
- Henderson, J. M., & Siefert, A. B. C. (2001). Types and tokens in transsaccadic object identification: Effects of spatial position and left–right orientation. *Psychonomic Bulletin and Review*, *8*, 753–760.
- Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 519–537.
- Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, *137*, 163–181.
- Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception and Psychophysics*, *69*, 172–184.
- Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 307–317.
- Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125–155). Cambridge, MA: MIT Press.
- Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention, and transsaccadic memory. *Visual Cognition*, *5*, 127–155.
- Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *41*–55.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.
- Kolers, P. A. (1972). *Aspects of motion perception*. Oxford: Pergamon Press.
- Kruschke, J. K., & Fragassi, M. M. (1996). The perception of causality: Feature binding in interacting objects. In *Proceedings of the eighteenth annual conference of the cognitive science society* (pp. 441–446). Hillsdale, NJ: Erlbaum.
- Leslie, A., Xu, F., Tremoulet, P., & Scholl, B. J. (1998). Indexing and the object concept: “What” and “where” in infancy. *Trends in Cognitive Sciences*, *2*, 10–18.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, underlie object persistence. *Psychonomic Bulletin and Review*, *14*, 1199–1204.
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and conquer: How object files adapt when a persisting object splits into two. *Psychological Science*, *15*, 420–425.
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition*, *96*, 67–92.
- Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science*, *15*, 866–871.
- Moore, C. M., Mordkoff, J. T., & Enns, J. T. (2007). The path of least persistence: Evidence of object-mediated visual updating. *Vision Research*, *47*, 1624–1630.
- Moore, M. K., Borton, R., & Darby, B. L. (1978). Visual tracking in young infants: Evidence for object identity or object permanence. *Journal of Experimental Child Psychology*, *25*, 183–198.
- Pylshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, *4*, 197–207.
- Pylshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, *11*, 801–822.
- Pylshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197.
- Richard, A. M., Luck, S. J., & Hollingworth, A. (2008). Establishing object correspondence across eye movements: Flexible use of spatiotemporal and surface feature information. *Cognition*, *109*, 66–88.
- Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind and Language*, *22*, 563–591.
- Scholl, B. J., & Pylshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.
- Scholl, B. J., Pylshyn, Z. W., & Franconeri, S. L. (1999). *When are spatiotemporal and featural properties encoded as a result of attentional allocation?* Paper presented at the conference of the Association for Research in Vision and Ophthalmology.
- Spivey, M., Richardson, D., & Fitneva, S. (2004). Thinking outside the brain: Spatial indices to linguistic and visual information. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 161–189). New York: Psychology Press.
- Tarr, M. J., Bulthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science*, *8*, 282–289.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, *40A*, 201–237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*, 111–153.
- Xu, F., Carey, S., & Welch, J. (1999). Infants' ability to use object kind information for object individuation. *Cognition*, *70*, 137–166.
- Zelinsky, G. J., & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception and Psychophysics*, *67*, 676–690.
- Zhang, W. W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235.