Establishing object correspondence across eye movements: Flexible use of spatiotemporal and surface feature information ☆

Ashleigh M. Richard [a,*], Steven J. Luck [b], Andrew Hollingworth [a]

[a] Department of Psychology, The University of Iowa, 11 Seashore Hall E, Iowa City, IA, 52242-1407, USA
[b] Center for Mind & Brain and Department of Psychology, University of California, Davis, USA

A R T I C L E   I N F O

A B S T R A C T

Visual input is frequently disrupted by eye movements, blinks, and occlusion. The visual system must be able to establish correspondence between objects visible before and after a disruption. Current theories hold that correspondence is established solely on the basis of spatiotemporal information, with no contribution from surface features. In five experiments, we tested the relative contributions of spatiotemporal and surface feature information in establishing object correspondence across saccades. Participants generated a saccade to one of two objects, and the objects were shifted during the saccade so that the eyes landed between them, requiring a corrective saccade to fixate the target. To correct gaze to the appropriate object, correspondence must be established between the remembered saccade target and the target visible after the saccade. Target position and surface feature consistency were manipulated. Contrary to existing theories, surface features and spatiotemporal information both contributed to object correspondence, and the relative weighting of the two sources of information was governed by the demands of the task. These data argue against a special role for spatiotemporal information in object correspondence, indicating instead that the visual system can flexibly use multiple sources of relevant information.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Visual perception seems effortless and seamless, yet the input for vision is frequently disrupted by events such as eye movements, blinks, and occlusion. Such disruptions create a *correspondence problem* that pervades natural vision. Consider the situation in which a person is cooking a meal and makes a saccadic eye movement from a toaster to a fork. Before the saccade, the toaster lies at the center of the retinal image, and the fork projects to a peripheral region of the retina. During the saccade, vision is suppressed (Matin, 1974), creating a gap in perceptual input. After the saccade, the fork lies near the center of the retinal image, and the toaster projects to the periphery (and, of course, the retinal positions of all other visible objects change as well). How does the visual system determine that the fork after the saccade is the same object that appeared in a different retinal location before the saccade, especially given that several objects might lie near the location of the fork? In other words, how is object correspondence established across the disruption and change introduced by the eye movement? A similar problem is generated by brief occlusion. If one of three running children briefly passes behind a parked car, how does the visual system determine that the child emerging from behind the car is the same child as the one that disappeared behind the car? These occlusion events may occur many times each day, and eye movements occur tens of thousands of times each day, with each one generating

the need to compute correspondence. Establishing object correspondence across perceptual disruption is therefore a fundamental problem the visual system must solve.

1.1. Object-file theory: object correspondence via spatiotemporal continuity

In an influential paper, Kahneman, Treisman, and Gibbs (1992) proposed a general solution to the problem of object correspondence, termed *object-file theory*. In this view, when an object is present in the visual field, a position marker, or *spatial index*, is assigned to the location occupied by the object. Kahneman et al. conceived of spatial indexes as contentless pointers to locations within an internal representation of external space (see also Pylyshyn, 2000). When an object is attended, perceptual properties of the object (color, shape, etc.) and other properties (identity, appropriate responses) are activated and become associated with, or bound to, the spatial index. This composite representation – the spatial index with associated object properties – forms the object file. Object files are proposed to be robust across brief disruptions in perceptual input, forming a short-term visual memory system that can be used to compute object correspondence.

A central claim of object-file theory is that because objects are addressed by their locations, object correspondence is computed on the basis of spatiotemporal continuity. A currently viewed object is treated as corresponding to a previously viewed object if the object's position over time is consistent with the interpretation of a continuous, persisting entity. For example, if a moving object passes behind an occluder, its object file will be maintained in memory during the occlusion. If the object re-appears in the appropriate spatial location at the appropriate time, it will be assigned the same index as it had been assigned before (Scholl & Pylyshyn, 1999). "Bobby" appearing from behind the parked car will be assigned the same index as had been assigned to "Bobby" disappearing behind the car, and object correspondence will have been established. Similarly, across a saccade, the relative positions of a few objects will be stored as object files, forming a spatial configuration. After the saccade, objects appearing in locations that match the locations in the remembered configuration will be treated as corresponding to the objects viewed in those locations before the saccade (Pylyshyn, 2000). If a fork that is perceived after the saccade occupies the same position within a configuration of objects that it did before the saccade, it will be assigned the same index as before the saccade, establishing object correspondence across the eye movement. In each of these cases, the object will be consciously perceived as a single, continuous entity.

Object-file theory further holds that the computation of object correspondence does not consult non-spatial properties of the object, such as shape, color, or meaning. This follows from the claim that the content of the file (e.g., surface feature and identity information) can be accessed only after spatial correspondence has been established (Kahneman et al., 1992). Therefore, surface feature information could not itself be used to establish correspondence. Moreover, if spatiotemporal information is consistent with the interpretation of a continuous object, object correspondence will be established even if surface feature and identity information are inconsistent with the interpretation of correspondence. This element of object-file theory allows an object to be perceived as continuous despite changes in surface features and identity, which Kahneman et al. argued is important given that the perceived properties of an object sometimes change. For example, if a running animal that one thought was a dog turns out to be a coyote, it will be perceived as a single, continuous entity despite the change in perceived identity.

The claim of position dominance in object-file theory might seem implausible on the face of it. One can think of many circumstances in which surface feature information is used to establish the correspondence between separate encounters with an object despite changes in object position. For example, if one leaves a colleague's office while the colleague is seated and then returns to find the colleague standing by a bookshelf in a different part of the office, one would have no difficulty determining on the basis of surface features (e.g., clothing, facial features) that the colleague is the same person who had been seated before. Kahneman et al. (1992; see also Treisman, 1992) proposed that this type of ability is supported by an object recognition mechanism that is to be distinguished from an object correspondence mechanism. According to this distinction, object recognition relies on long-term memory (LTM) for the visual form of objects that is largely position invariant. And object recognition does not lead to the perceptual experience of seeing an object as continuously present. In contrast, object correspondence across brief disruptions (such as a saccade) depends on the short-term maintenance of position-specific object files, is governed by spatiotemporal continuity, and generates the experience of object persistence.

The central evidence that object correspondence consults spatial, but not surface feature, information is derived from the original object reviewing paradigm developed by Kahneman et al. (1992). In the standard version of this task, 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 test letter was present in the preview display. In addition, there is a further benefit when the letter appears in the same object as it had appeared originally. This "same-object benefit" has been interpreted as reflecting facilitation when the test letter matches the contents of the object file associated with the new location of the object (and thus it is more accurately described as a "consistent position benefit"). The benefit is observed when there is spatiotemporal information consistent with an object's persistence as a single entity, such as in a sequence of apparent-motion frames (Kahneman et al., 1992).

Critically, however, surface feature information has been found to be ineffective in establishing object correspondence in this paradigm. In Kahneman et al. (1992, Experiment 6), participants were presented with two boxes above and below fixation, and the letters that appeared in the boxes were drawn in different colors. The

letters and boxes disappeared, and then the two boxes were presented to the left and right of fixation. Spatial correspondence was ambiguous. A letter appeared in one of the two boxes, and color consistency of the target letter was manipulated. Participants were no faster to name the letter 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 and test display. Mitroff and Alvarez (2007) replicated and extended this finding in a modified object-reviewing paradigm. Letters appeared in objects that differed on a number of salient surface feature properties (color, shape, size, luminance, topography, and polarity). Letter retrieval from memory was found to be independent of the association between the letter and the particular object (defined by its surface features). These findings appear to imply that surface feature information plays no role in establishing object correspondence, at least within the object-reviewing paradigm.

The apparent dominance of spatiotemporal information in object correspondence has been important in the translation of the object-file framework to other domains of perception and cognition. The multiple object tracking (MOT) paradigm (Pylyshyn & Storm, 1988; Scholl, Pylyshyn, & Feldman, 2001) has been used to examine spatial mechanisms of object correspondence under conditions of object motion. In these tasks, participants track a subset of identical objects as the objects move unpredictably. Because the objects are identical, they can be individuated only by virtue of their spatial position over time. MOT studies have shown that spatial information is sufficient to track as many as 4–5 moving objects simultaneously, and that tracking is robust across brief occlusion (Scholl & Pylyshyn, 1999). In addition, tracking is robust even if the surface features of tracked objects change (Scholl, Pylyshyn, & Franconeri, 1999), and participants can track object locations without keeping track of the surface properties of those objects (Horowitz et al., 2007; Pylyshyn, 2004). The object-file framework has also played an important role in understanding perceptual development in infants. 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 (Feigenson & Carey, 2005; Van de Walle, Carey, & Prevor, 2000; Xu, Carey, & Welch, 1999).

1.2. The information used to establish object correspondence across saccades

The object-file framework also has had a significant impact on theories of object memory and correspondence across eye movements. Irwin and colleagues (Irwin, 1992; Irwin & Andrews, 1996; Irwin & Gordon, 1998) proposed that the object representations maintained across an eye movement are object files, with perceptual properties of remembered objects bound to spatial locations (see also Zelinsky & Loschky, 2005). In addition, Henderson and colleagues (Henderson, 1994; Henderson & Anes, 1994; Henderson & Siefert, 2001) replicated the Kahneman et al. position consistency effects across saccades, suggesting

that object correspondence across saccades depends on the same mechanisms used to compute correspondence across object motion.

However, researchers studying correspondence across saccades have not been strongly committed to the idea that correspondence operations consult only spatiotemporal information. Currie, McConkie, Carlson-Radvansky, and Irwin (2000) proposed that before a saccade, an object is selected as the saccade target. The allocation of attention to the saccade target leads to the consolidation of information from that object into visual short-term memory (VSTM) (Hollingworth, Richard, & Luck, 2008; Irwin & Gordon, 1998). The encoded object information could include both spatial and non-spatial features of the saccade target (such as color, shape, or identity). These features are maintained in VSTM across the saccade. After the saccade, the visual system compares objects lying near the saccade landing position with the stored features of the target object. If an object matching the remembered features of the target is found near the saccade landing position, correspondence is established. Although this proposal assumes that surface feature information could be used to map objects across saccades, no empirical tests were conducted to specifically determine whether surface feature information, as opposed to spatiotemporal information, is used in this manner.

Recently, Hollingworth et al. (2008) provided the first direct evidence that surface feature information can drive the computation of object correspondence across saccades. The theoretical assumptions and method of the present study draw from those of Hollingworth et al. (2008), and we will therefore review them in some detail.

Although each eye movement generates the need to compute object correspondence, the problem of correspondence is particularly pressing when the eyes fail to land on the target of the saccade. Saccades are notoriously prone to error, with saccade errors occurring on as many as 40% of trials in laboratory studies (for a review, see Becker, 1991). When the eyes miss the saccade target, there is typically a short fixation followed by a corrective saccade to bring the target object onto the fovea (e.g., Deubel, Wolf, & Hauske, 1982). To correct gaze to the appropriate object, correspondence must be established between the target object visible before the saccade and the target object visible after the saccade. Such an operation is trivial if there is only one visible object, as in most laboratory studies of gaze correction. But in the real world, there will often be multiple objects near the saccade landing position. If the saccade from the toaster to the fork misses the fork, the fork must be found among other nearby objects (spoon, measuring cup, etc.). Because there are multiple visible objects that could potentially be the target, efficient correction to the original saccade target requires memory for that object across the saccade. Specifically, Hollingworth et al. (2008) hypothesized that VSTM is used to remember saccade target properties across the eye movement so that after an inaccurate saccade, the target can be discriminated from other nearby objects and gaze efficiently corrected.

In the Hollingworth et al. gaze correction paradigm (illustrated in Fig. 1), participants fixated the center of an array of differently colored objects. One object was cued,
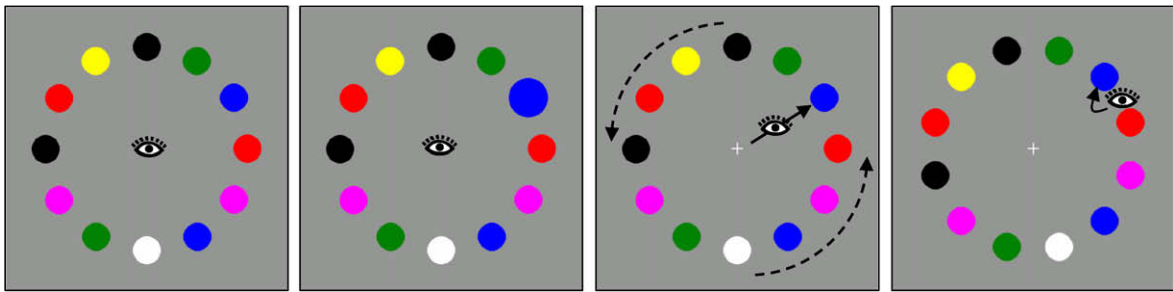
Fig. 1. Sequence of events in a rotation trial of Hollingworth et al. (2008). During the saccade to the cued object, the array was rotated 1/2 object position (counter clockwise in this sample trial), causing the eyes to land between the target object and an adjacent distractor. Memory for the target color was required to correct gaze to the appropriate object. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

and participants executed a saccade to this object. On some of the trials, the array was rotated 1/2 object position clockwise or counterclockwise during the saccade, causing the eyes to land midway between the target object and a distractor (the neighboring object), which simulated a saccade error. The rotation could not be perceived directly, because vision is suppressed during saccades (Matin, 1974). In this method, spatial position was completely uninformative for the purpose of computing object correspondence. The only means to discriminate the target from distractor, and execute the appropriate corrective saccade, was to remember a surface feature property of the target (its color) and compare this information with objects visible after the saccade. The accuracy and speed of gaze correction in this full-array condition was compared with a single-object control condition that presented only one object, the target, and therefore did not require memory to correct gaze.

The principal finding was that gaze correction in the full-array condition was highly accurate and efficient. Participants corrected gaze to the appropriate target object on 98% of trials, and the mean latency of these corrections (240 ms) was only 39 ms longer than for corrections in the single-object control condition, which did not require memory. In addition, we can be confident that correspondence in this paradigm is established on the basis of visual memory for the saccade target, rather than on the basis of a verbal description: a secondary VSTM load significantly impaired gaze correction, but a secondary verbal memory load did not. Finally, VSTM-based gaze corrections were largely unconscious and automatic. Gaze was often corrected to the remembered saccade target despite explicit instructions to avoid doing so. Together, these data indicate that memory for an object's visual surface features can be used efficiently to establish correspondence between the saccade target object visible before and after a saccade, a situation that arises thousands of times each day in naturally occurring visual perception.

Although the Hollingworth et al. (2008) results demonstrate that surface feature information can be used to establish object correspondence across saccades, this may reflect a special case in which spatial position was noninformative. That is, surface features may be consulted only when correspondence cannot be established on the basis of spatiotemporal information. Thus, spatial information might be the sole determinant of correspondence

across saccades when spatial information is informative, consistent with the object-file theory assumption of position dominance. One of the central goals of the present study was to examine the relative contribution of spatial and surface feature information to transsaccadic correspondence when both types of information were informative.

2. The present study

Given that the problem of object correspondence is most frequently caused by eye movements (saccades occur much more frequently than any other type of perceptual disruption), a general theory of object correspondence must account for correspondence across saccades. Thus, we tested the object-file theory claim of spatiotemporal correspondence within a gaze correction task that required mapping the pre-saccade target object to the post-saccade target object. The critical manipulation created a conflict between spatiotemporal information and surface feature information. This allowed us to test a range of hypotheses regarding the relative weighting of the two sources of information.

The basic paradigm used in the present study is illustrated in Fig. 2. Participants saw two objects arrayed in a column to the right of fixation. There was a gap between the two objects, roughly the size of an additional object. One of the two objects was cued as the saccade target, and the participant executed an eye movement to that object. On 2/5 of the trials, the array shifted either up or down during the saccade to the target, when vision is suppressed. The shift was half of the distance between the two objects, typically causing the eyes to land midway between them. For example, if the top object was cued and a saccade was initiated toward that object, the two objects would shift up, causing the eyes to land between the two objects rather than on the saccade target. Alternatively, when the bottom object was cued and a saccade initiated to that object, the two objects would shift down, again causing the eyes to land in the gap between the two objects. When the array shifted, a corrective saccade was required to bring gaze to the target object. Because the shift occurred during the saccade, correction of gaze to the appropriate object required information from the pre-saccade array (such as the color or relative location of the target object) to be stored across the saccade and compared
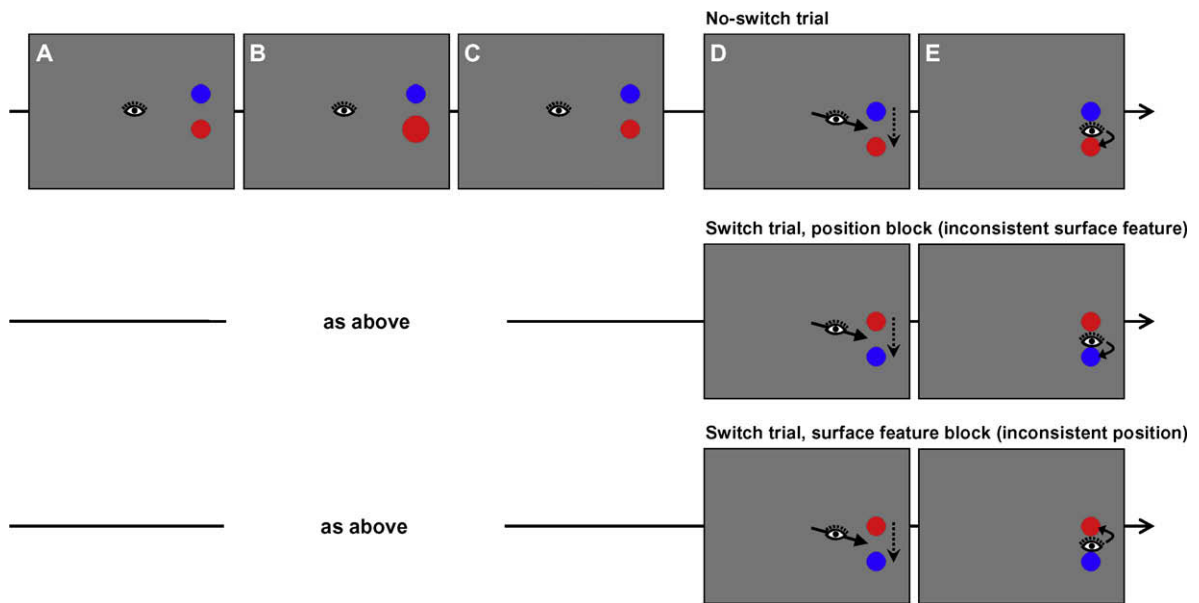
Fig. 2. Sequence of events in a "shift" trial for each of the three principal conditions in Experiment 1. In these examples, the array objects were shifted down 1/2 object position during the saccade to the target, causing the eyes to land between the two objects. The top row shows a no-switch trial in which the two objects retained their original properties across the saccade. The middle and bottom rows show switch trials, in which the two objects traded properties across the saccade. The middle row shows a switch trial in the position block. Participants corrected gaze to the object in the correct target position despite inconsistent color at that location. The bottom row shows a switch trial in the surface feature block. Participants corrected gaze to the object with the correct target color despite inconsistent position information. In the grayscale version of this figure, the two colors (red and blue) are represented by white and black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

with perceptual information available after the saccade. Although this is a somewhat artificial situation that was designed to provide experimental control over saccade errors, it is important to stress that it reflects a common real-world scenario, in which the eyes do not land on the intended target and several items lie near fixation after the saccade.

In this paradigm, correspondence could be established either on the basis of spatial information (the relative position of the target) or on surface feature information (the color of the target). To examine the relative contribution of the two sources of information, on half of the shift trials (1/5 of all trials) the two objects switched properties during the saccade to the cued object. Consider the sample trials from Experiment 1 depicted in Fig. 2 (middle and bottom rows). During the saccade to the target, the objects were shifted, and in addition, the colors switched relative locations. This placed position information and surface feature information at odds. Spatial information (that the cued object was in the bottom position) would map the object in the lower position to the original saccade target. Surface feature information (that the cued object was red, depicted as white in the grayscale version of Fig. 2) would map the object in the upper position to the original saccade target. Placing the two sources of information in conflict allowed us to examine which source of information was given more weight in establishing correspondence across the saccade.

In our experiments, the target was defined by its relative position in one block of trials (*position block*) and by its surface features in another block (*surface feature block*).

In the position block, for example, a brief looming of a red item in the lower position would indicate that the participant should fixate the lower object irrespective of its color. In the surface feature block, a brief looming of this same object would indicate that the participant should fixate the red item irrespective of its spatial position. On the critical switch trials, participants had to correct gaze to the appropriate target despite inconsistent evidence on the other dimension. In the position block, participants would direct their eyes to the object in the appropriate relative location despite the fact that its color was inconsistent with the original target color (Fig. 2, middle row). In the surface feature block, participants would direct their eyes to the appropriately colored object despite the fact that its location was inconsistent with the original target location (Fig. 2, bottom row). On no-switch trials, the objects did not switch properties, providing a baseline measure of gaze correction efficiency when the two sources of information were consistent.

Two dependent measures provided data regarding the efficiency of object correspondence and gaze correction. Gaze correction accuracy was the proportion of trials on which the eyes were directed first to the appropriate target object after landing between target and distractor. Gaze correction latency was the duration of the fixation before the corrective saccade when a single corrective saccade was directed to the appropriate object (similar to correct RT).

The gaze correction paradigm provides a robust, online measure of correspondence operations across very brief gaps in perceptual input (30–50 ms saccades), allowing us

to probe the moment-to-moment use of memory to map objects across perceptual disruption (Hollingworth et al., 2008). The gaze correction method is to be preferred, in a number of ways, over the original object-reviewing paradigm (Kahneman et al., 1992). First, the original object-reviewing paradigm measures correspondence indirectly by examining perception of an object (a letter) associated with a different object whose continuity is manipulated (a box). The gaze correction paradigm tests correspondence directly by manipulating features of the saccade target object itself. Second, in the object-reviewing paradigm, evidence that the visual system has treated an object as continuous across change is derived from an indirect assessment of perceptual priming obtained well after the object change has occurred. In contrast, the gaze correction paradigm examines object continuity by directly measuring the moment-by-moment allocation of overt attention to objects, providing a real-time window on the correspondence operations that serve to map objects across brief perceptual disruption. Third, the gaze correction paradigm simulates a real-world behavior for which correspondence is critically important (gaze correction), whereas the mapping of the object-reviewing paradigm onto real-world behavior is not as direct. Finally, the object-reviewing paradigm provides very strong spatiotemporal cues, in the form of continuously moving squares, but it does not naturally lend itself to strong surface feature cues. In contrast, the present paradigm can put surface features and location on an equal footing by using the same stimuli and simply varying which dimension is task-relevant (cf. the middle and bottom rows of Fig. 2).

We tested three hypotheses regarding the possible roles of spatial and surface feature information in the computation of transsaccadic object correspondence:

1. Only spatial information is consulted. This *strong spatiotemporal hypothesis* proposes that spatiotemporal information is the only determinant of object correspondence (Kahneman et al., 1992; Mitroff & Alvarez, 2007). In its strongest form, there would be no exceptions to this rule. However, we have already demonstrated that surface features can be used when spatial information is non-informative (Hollingworth et al., 2008). Thus, a modified version of this hypothesis might state that surface feature information is consulted when it is the only available cue to correspondence. In the present experiments, however, spatial information was highly informative. Spatial position was associated with the target object on 80% of trials in the surface feature block and 100% of trials in the position block. The strong spatiotemporal hypothesis predicts that gaze correction should be more efficient when participants correct gaze on the basis of position than when they correct on the basis of surface features. Most importantly, because surface features are proposed to play no role in object correspondence when position is informative, inconsistent surface feature information (in the switch trials of the position block) should not impair gaze correction performance. Finally, inconsistent position information (in the switch trials of the surface feature block) should generate significant interference with gaze cor-rection, slowing corrections and possibly causing participants to incorrectly direct gaze to the object in the original relative position rather than to the object with the original color.

2. Both spatiotemporal and surface feature information are consulted, but position dominates. This *weak spatiotemporal hypothesis* assumes that spatiotemporal information is the primary determinant of object correspondence, but surface feature information is consulted as well. If so, then in the surface feature block, inconsistent position should impair gaze correction. And in the position block, inconsistent surface features should also impair gaze correction. Critically, the interference generated by inconsistent position should be larger than the interference generated by inconsistent surface features.

3. Both are consulted, and the relevant information is consulted flexibly on the basis of task demands. This *flexible correspondence hypothesis* holds that if surface feature information is made salient within the task, it will dominate correspondence across the saccade. If spatial information is made salient, it will dominate instead. To examine this possibility, we manipulated the means by which the saccade target was selected initially. If the saccade target is selected by virtue of its spatial position (e.g., the participant is cued to generate a saccade to the top object), then spatial position will dominate in computing correspondence. However, if the saccade target is selected by virtue of surface features (e.g., the participant is cued to generate a saccade to the red object), surface feature information will play the primary role in establishing object correspondence.

These hypotheses were tested in five experiments. The results favored the flexible correspondence hypothesis. When the saccade cue created a spatially local signal at the target position (Experiments 1 and 2) or when the saccade target was selected on the basis of its relative position (Experiment 5), correspondence was dominated by position consistency. However, when the target object was selected by virtue of its surface features (Experiments 3 and 4), correspondence was dominated by surface feature consistency.

3. Experiment 1

In Experiment 1, we examined gaze correction using an array of two color disks, as illustrated in Fig. 2. The target object was cued by the rapid expansion and contraction of that object, which generated a spatially local signal at the target position.

3.1. Method

3.1.1. Participants
Sixteen participants from the University of Iowa community completed the experiment. They received course credit or were paid. All participants reported normal vision.

3.1.2. Stimuli and apparatus
Arrays consisted of two color disks, one red and one blue, randomly assigned to the two possible locations.

The two disks were presented to the right of fixation, above and below the horizontal midline, and were equidistant from fixation. Disks subtended 1.6° and were centered 5.9° from central fixation. The distance between the centers of the disks was 3.0°. A sample display is shown in Fig. 2. The saccade target cue was the rapid expansion of the target object to 140% of its original size and contraction back to the original size over 50 ms of animation. On shift trials, the two objects were shifted together 1.5° vertically, either up or down, during the saccade.

Stimuli were displayed on a 17-in. CRT monitor with a 120 Hz refresh rate. Eye position was monitored by a video-based, ISCAN ETL-400 eyetracker sampling at 240 Hz. A chin and forehead rest was used to maintain a constant viewing distance of 70 cm and to minimize head movement. The experiment was controlled by E-prime software. Gaze position samples were transferred in real time from the eyetracker to the computer running E-prime. E-prime then used gaze position data to control trial events (such as transsaccadic shift) and saved the raw position data to a file that mapped eye events and stimulus events.

Array shift during the saccade to the target was accomplished using a boundary technique. An invisible, vertical boundary was defined 3.2° to the right of central fixation. After the saccade target cue, the computer monitored for an eye position sample to the right of the boundary, and on array shift trials, the shifted array was then written to the screen. Pilot testing ensured that the shift was completed before the eyes landed, so that the change occurred during the saccade itself and could not be perceived directly (see Hollingworth et al., 2008).

3.1.3. Procedure

All trials began with participants fixating a cross at screen center. Eyetracker calibration was checked, and the participant was recalibrated as needed. The experimenter then initiated the trial. After a delay of 300 ms, the two objects were displayed for 1000 ms. One of the two objects was cued (by expansion and contraction). Participants were instructed to direct their eyes to the cued object as quickly as possible. Once the appropriate object was fixated, a box appeared around that object for 400 ms, and the trial ended. Thus, participants received feedback on every trial, because the trial did not advance until the target was fixated.

On 3/5 of the trials, the object array did not change during the eye movement (no shift trials). On the remaining 2/5 of trials, the two disks shifted up (if the top disk was cued) or down (if the bottom disk was cued) one half object position. The eyes typically landed between the disks, and a corrective saccade was made to the target disk. On half of these shift trials, the two objects switched properties.

In the surface feature block, participants were instructed to always direct the eyes to the object that had the same color as the cued object, regardless of its position. In the position block, participants were instructed to always direct the eyes to the object that occupied the same relative position (top or bottom) as the cued object, regardless of its color. Participants were informed of possible array changes (shift and property switch) and the need to redirect their eyes on some trials.

In each of the two blocks, participants completed 10 practice trials followed by 120 experiment trials. Seventy-two of the experiment trials (60%) were no-shift trials in which the array did not change across the saccade. The top object was cued on half of the trials and the bottom object was cued on the other half. On the remaining 48 trials (40%) the objects shifted, upward or downward with equal probability. In addition, on half of the shift trials, the two objects switched properties. Block order was counterbalanced across participants. The entire experiment session lasted approximately 45 min.

3.1.4. Data analysis

Eye tracking data analysis was conducted offline using dedicated software. A velocity criterion (eye rotation >31°/s) was used to define saccades. During a fixation, the eyes are not perfectly still. Fixation position was calculated as the mean position during a fixation period weighted by the proportion of time at each sub-location within the fixation. These data were then analyzed with respect to scoring regions surrounding each of the two post-saccade objects. Object scoring regions were circular and had a diameter of 1.9°, 20% larger than the color disks themselves.

Array-shift trials were eliminated from analysis if the eyes initially landed on an object rather than between objects, if more than one saccade was required to bring the eyes from central fixation to the general region of the object array, if the eyetracker lost track of the eye, or if corrective saccade latency was greater than 500 ms. The large majority of eliminated trials were those in which the eyes landed on an object, reflecting the fact that saccades are often inaccurate. A total of 27.8% of trials was eliminated. For all experiments, trial elimination affected neither the data pattern nor the statistical significance of any analysis. For all trials in Experiment 1 (shift and no-shift), the mean latency of the initial saccade from central fixation to the target object was 275 ms.

3.2. Results

The array-shift trials were of central interest for examining object correspondence across saccades. No-shift trials served as filler trials, and thus were not included in the analyses. Correction accuracy (the percentage of trials on which the target was the first disk fixated after landing between target and distractor) and correction latency (the duration of the fixation before an accurate corrective saccade) were examined for array-shift trials in each of the conditions created by the 2 (surface feature block, position block) × 2 (no switch, property switch) design. The data are reported in Fig. 3.

Gaze correction on no-switch trials in both blocks was highly accurate, with 99.6% correct in the position block and 99.7% correct in the surface feature block, $t(15) = 1.00$, $p = .33$. Gaze correction was also highly efficient, with mean latencies of 229 ms in the position block and 229 ms in the surface feature block, $t(15) = 0.02$, $p = .98$. Thus, participants were able to reliably and quickly
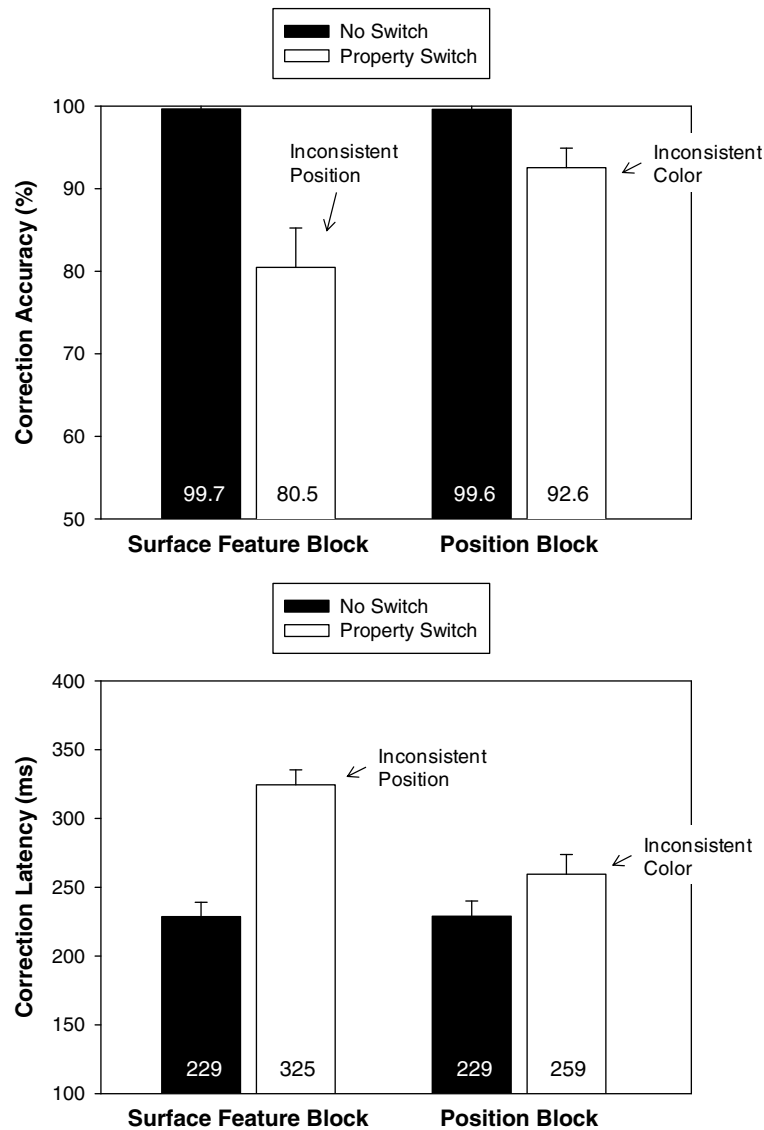
Fig. 3. Experiment 1. Mean gaze correction accuracy (top) and latency (bottom) as a function of the relevant dimension for gaze correction (position block, surface feature block) and property switch. Error bars are standard errors of the means.

establish object correspondence when both spatial and surface feature information specified the target object after the saccade.

We next compared the property-switch and no-switch conditions to examine the interference generated by inconsistent surface feature information or by inconsistent position information. In the surface feature block, mean correction accuracy on property-switch trials was significantly lower (80.5%) than accuracy on the no-switch trials (99.7%), $t(15) = 4.09$, $p < .001$. In addition, mean correction latency was significantly longer for property-switch trials (325 ms) than for no-switch trials (229 ms), $t(15) = 12.02$, $p < .001$. Thus, participants corrected gaze to the item with the appropriate color on a high proportion of trials even when its relative location within the array had changed, indicating that they could use surface features to establish

correspondence between the pre-saccade and post-saccade target object. However, position inconsistency caused gaze correction to be slower and less accurate.

In the position block, mean correction accuracy on property-switch trials was reliably lower (92.6%) than accuracy on the no-switch trials (99.6%), $t(15) = 3.13$, $p < .01$, and mean correction latency was significantly longer on property-switch trials (259 ms) than on no-switch trials (229 ms), $t(15) = 2.15$, $p < .05$. Thus, although the participants could make gaze corrections to the item in the appropriate relative location even when the color of the item changed, color inconsistency generated significant interference.

Although inconsistency produced interference for both conditions, inconsistent spatial information generated greater interference than inconsistent surface feature
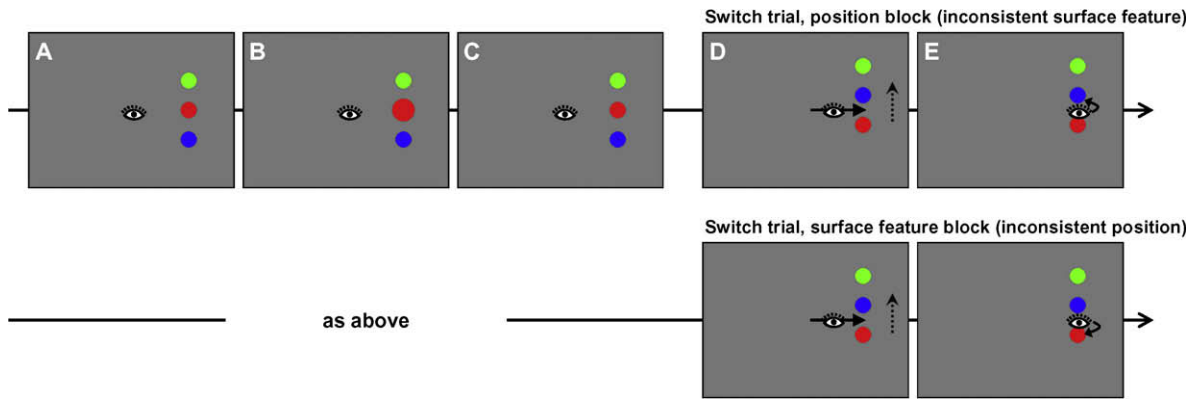
Fig. 4. Sequence of events in the property-switch trials of the Experiment 1 control study. Top row: position block. Bottom row: surface feature block. In the grayscale version of this figure, the three colors (blue, red, and green) are represented by black, white, and striped. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

information. For gaze correction accuracy, the size of the effect of inconsistent spatial information in the surface feature block (19.2%) was reliably larger than the effect of inconsistent surface feature information in the position block (7.1%), $t(15) = 2.20$, $p < .05$. For gaze correction latency, the effect of inconsistent spatial information in the surface feature block (96 ms) was reliably larger than the effect of inconsistent surface feature information in the position block (31 ms), $t(15) = 5.00$, $p < .001$.

One possible concern with the Experiment 1 method is that the cued object in the position block always predicted the direction of the possible corrective saccade. If the top object was cued and a shift occurred, the shift was always upward (requiring an upward correction), and if the bottom object was cued and a shift occurred, it was always downward (requiring a downward correction). Participants might have used this information to pre-program a corrective saccade in the direction of the possible shift. Such pre-programming could have artificially reduced an effect of surface feature match. We conducted a control experiment to eliminate the possibility that corrective saccades were pre-programmed. Displays consisted of three objects, instead of two, arranged in a vertical column (Fig. 4). One object was green, one blue, and one red, randomly assigned to the three locations. The center object was always cued, and the array could shift either up or down during the saccade, causing the eyes to land between the center object and either the top or bottom object. On property-switch trials, the center object and one other object (the top object if the array shifted down; the bottom object if the array shifted up) traded properties. Because the center object was always cued, participants could not predict the direction of the shift and could not pre-program a corrective saccade. In all other respects, the control experiment was identical to Experiment 1.

The control results replicated all features of the main Experiment 1 results. In the surface feature block, mean correction accuracy on property-switch trials (84.2%) was reliably lower than accuracy on no-switch trials (99.7%), $t(15) = 4.58$, $p < .001$, and correction latency was significantly longer for property-switch trials (287 ms) than for no-switch trials (196 ms), $t(15) = 7.23$, $p < .001$. In the posi-

tion block, correction accuracy on property-switch trials (91.2%) was reliably lower than accuracy on no-switch trials (99.7%), $t(15) = 6.58$, $p < .001$, and correction latency was significantly longer for property-switch trials (247 ms) than for no-switch trials (210 ms), $t(15) = 4.96$, $p < .001$. Finally, the magnitude of interference introduced by inconsistent spatial position was larger than the magnitude of interference generated by inconsistent surface features, both for correction accuracy, $t(15) = 1.83$, $p = .086$, and latency, $t(15) = 5.53$, $p < .001$.

3.3. Discussion

Experiment 1 generated four principal findings. First, gaze correction was, in general, highly accurate and efficient. In the no-switch trials, correction accuracy was essentially perfect, and mean correction latency was 229 ms. (In the control experiment, mean correction latency for no-switch trials was 201 ms). These latencies fell within the range of latencies observed for gaze corrections in Hollingworth et al. (2008), which were found to be largely independent of participant awareness and strategic control. Thus, we can be confident that the present paradigm probed the rapid, online operations used to establish object correspondence across saccades.[1]

Second, gaze correction to a color-defined target was slower and less accurate when the spatial position of the target object changed between the pre-saccade array and the post-saccade array. This result is an analog of the consistent-position benefit observed in earlier object-file stud-

---

[1] It is useful to compare these saccadic reaction times with manual reaction times. Even for the detection of a simple luminance target, manual responses typically have a mean latency of 250 ms or longer (e.g., Posner, Snyder, & Davidson, 1980). Our mean latency of 229 ms in the no-shift trials was observed despite the multiple component operations required after the eyes landed: (1) comparison of the remembered target properties with objects visible near the landing position; (2) selection of the object that best matches the remembered properties; and (3) computation and initiation of a saccade to that object. Our results are consistent with recent findings of exceedingly fast saccadic reaction times in studies requiring object discrimination (Kirchner & Thorpe, 2006) and visual change detection (Hyun, Woodman, Vogel, Hollingworth, & Luck, in press).
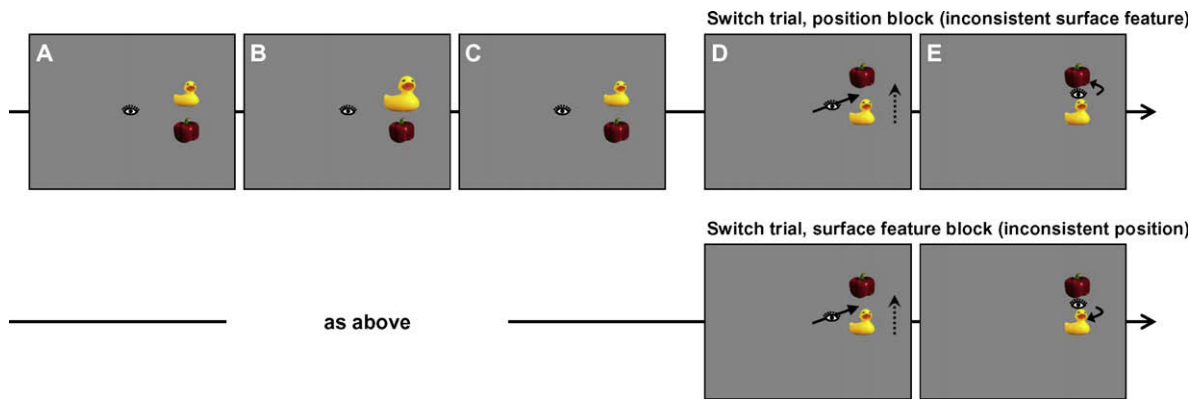
Fig. 5. Sequence of events in the property-switch trials of Experiment 2. Top row: position block. Bottom row: surface feature block.

ies (Henderson, 1994; Kahneman et al., 1992; Mitroff & Alvarez, 2007) and demonstrates that the present gaze correction paradigm is sensitive to position consistency effects.

Third, gaze correction to a location-defined target was slower and less accurate when the color of the target object changed between the pre-saccade array and the post-saccade array. This effect of surface feature consistency provides direct evidence that surface feature information is used to establish object correspondence across brief visual disruptions. It stands in contrast with traditional object-file experiments that have found no influence of surface feature consistency (Kahneman et al., 1992; Mitroff & Alvarez, 2007). The surface feature effect cannot be attributed to a special case in which position information is non-informative (as in Hollingworth et al., 2008), because remembered position perfectly predicted target location after the saccade in the position block. Indeed, the optimal strategy for the position block would have been to use only spatiotemporal information and to disregard surface features, and yet the participants were unable to do so. The results of this experiment therefore falsify the strong spatiotemporal hypothesis within the domain of transsaccadic correspondence. The strong spatiotemporal hypothesis is a central assumption of object-file theory, and object-file theory cannot accommodate the present findings without modification of its basic tenets about how object information is addressed and accessed within visual memory. These issues will be discussed in Section 8.

Finally, the effects of position consistency were reliably larger than the effects of surface-feature consistency. Thus, position information was weighted more heavily than surface feature information in Experiment 1. This finding provides initial support for the weak spatiotemporal hypothesis, which claims that although position and surface feature information contribute to object correspondence, position information is given precedence. However, it is difficult to control the magnitude of a particular manipulation across different perceptual dimensions such as color and space. It is possible that the position differences between the two objects in this experiment were simply more salient than the surface feature differences. In Experiment 2, the method was modified to determine

whether greater weighting of position would still be observed when surface feature differences between the two objects were made as salient as possible.

4. Experiment 2

Experiment 2 examined gaze correction to real-world, complex objects (see Fig. 5). The objects differed from each other on many dimensions that are potentially available for encoding into transsaccadic VSTM (e.g., shape, color, texture). In addition, the objects differed in their semantic category, providing a further non-spatial cue to object correspondence. By maximizing the non-spatiotemporal differences between objects, Experiment 2 provided a strong test of the position-primacy claim of the weak spatiotemporal hypothesis.

4.1. Method

4.1.1. Participants

Sixteen new participants from the University of Iowa community completed the experiment. They either received course credit or were paid. All participants reported normal vision.

4.1.2. Stimuli and apparatus

The stimuli and apparatus were the same as in Experiment 1 with the following exceptions. Arrays consisted of two color photographs of real-world objects. The two objects used on a given trial were selected at random without replacement from 48 possible objects. All 48 objects were artifacts and differed at the basic level of categorization (see examples in Fig. 5). Object photographs were obtained from the Hemera database and were resized to fit within a $3.2° \times 3.2°$ region. Objects were centered 7.5° from fixation, and the distance between the centers of the objects was 5.8°.

4.1.3. Procedure

The procedure was identical to that of Experiment 1. In the position block, participants were instructed to direct gaze to the object appearing in the cued position, regardless of which object appeared in that position. In the

surface feature block, participants were instructed to direct gaze to the cued object (e.g., to direct their eyes to the "doll" if the doll was cued), regardless of the position that object occupied.

4.1.4. Data analysis

Eyetracking data were analyzed in the same manner as in Experiment 1. Object scoring regions were square, 3.2° × 3.2°. A total of 34% of the array-shift trials was eliminated for the reasons specified in Experiment 1, but this had no impact on the pattern of results. For all trials (shift and no-shift), the mean latency of the initial saccade from central fixation to the target object was 237 ms.

4.2. Results

The results are shown in Fig. 6. Gaze correction on no-switch trials was highly accurate, with 99.4% correct in the position block and 100.0% correct in the surface feature block, $t(15) = 1.46$, $p = .17$. Mean correction latency on no-shift trials was 186 ms in the position block and 193 ms in the surface feature block, $t(15) = 0.80$, $p = .44$. Again, participants efficiently established object correspondence when both position and surface feature information specified the target. Note that gaze correction latency to real-world objects in Experiment 2 (∼190 ms) was considerably faster than correction to color disks in Experiment 1 (∼230 ms). It is difficult to compare these
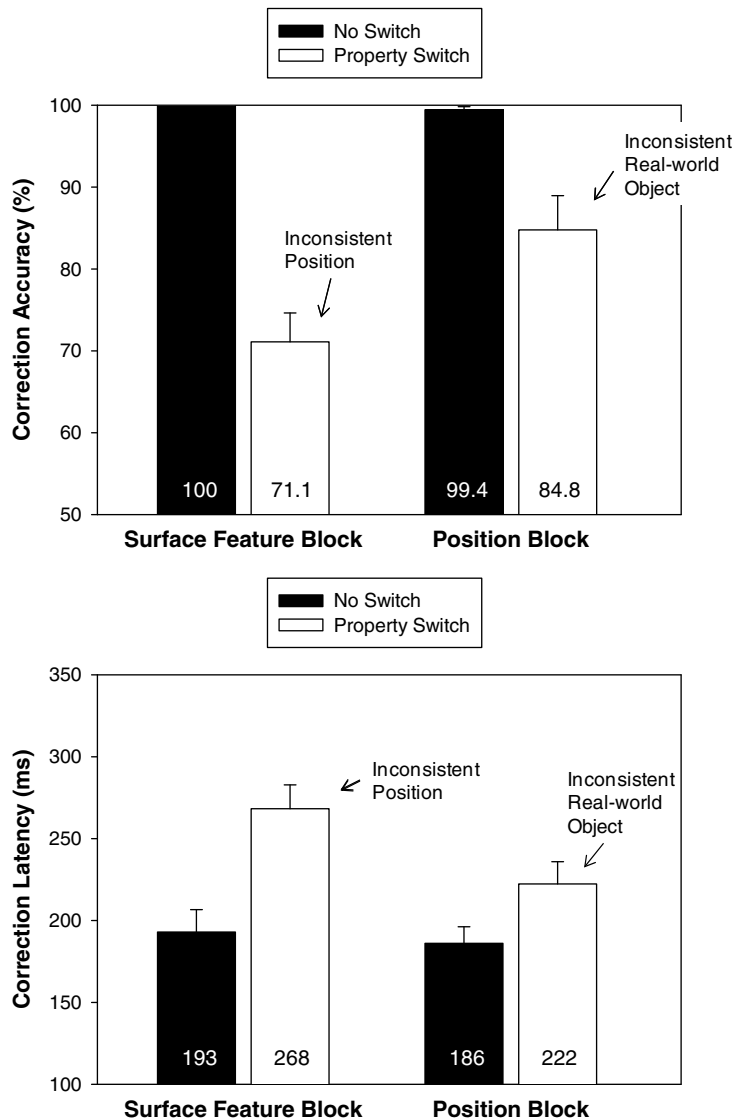

Fig. 6. Experiment 2. Mean gaze correction accuracy (top) and latency (bottom) as a function of the relevant dimension for gaze correction (position block, surface feature block) and property switch. Error bars are standard errors of the means.

latencies directly, however, because the distance between the two objects was greater in Experiment 2 than in Experiment 1, and corrective saccade latency is inversely related to the distance of the correction (Deubel et al., 1982). Thus, the latency difference could have been caused simply by differences in the distance of correction. Nevertheless, saccadic RTs of less than 200 ms in Experiment 2 indicate an extraordinarily efficient use of memory to locate the target object after the saccade, consistent with the assumption that VSTM supports (and the gaze correction paradigm probes) the online mapping of objects across saccades (Currie et al., 2000; Hollingworth et al., 2008).

In the surface feature block, mean correction accuracy on property-switch trials (71.1%) was reliably lower than accuracy on no-switch trials (100.0%), $t(15) = 8.20$, $p < .001$. And mean correction latency was significantly longer for property-switch trials (268 ms) than for no-switch trials (193 ms), $t(15) = 6.13$, $p < .001$. In the position block, mean correction accuracy on property-switch trials (84.8%) was reliably lower than accuracy on no-switch trials (99.4%), $t(15) = 3.49$, $p < .005$. And mean correction latency was significantly longer for property-switch trials (222 ms) than for no-switch trials (186 ms), $t(15) = 4.59$, $p < .001$.

As in Experiment 1, the magnitude of interference introduced by inconsistent spatial position was larger than the magnitude of interference generated by inconsistent surface features. For gaze correction accuracy, the size of the effect of inconsistent spatial information in the surface feature block (28.9%) was reliably larger than the size of the effect of inconsistent surface feature information in the position block (14.7%), $t(15) = 2.26$, $p < .05$. For gaze correction latency, the effect of inconsistent spatial information in the surface feature block (75 ms) was reliably larger than the effect of inconsistent surface feature information in the position block (36 ms), $t(15) = 2.84$, $p < .05$.

4.3. Discussion

Experiment 2 replicated the main findings of Experiment 1 with complex real-world objects that maximized non-spatiotemporal differences between objects (including both surface features and meaning). Although both position and surface feature information were consulted in object correspondence and gaze correction, the interference generated by inconsistent position was significantly larger than that generated by inconsistent surface features. In particular, inconsistent position information generated correction errors on almost 30% of trials. On these trials, participants were instructed, for example, to correct gaze to the "doll", and yet they corrected gaze to the object in the remembered target position despite the fact that the object was not a doll. This quite remarkable dependence on spatial information in Experiment 2 is consistent with the weak spatiotemporal hypothesis.

The pattern of results is not, however, consistent with the strong spatiotemporal hypothesis, because gaze corrections in the position block were slower and less accurate when the objects at the two locations were swapped. Thus, surface feature information must have been encoded into working memory and bound with relative location information; otherwise there could have been no effect of

swapping the relative locations of the objects (see Johnson, Hollingworth, & Luck, 2008; Wheeler & Treisman, 2002). Moreover, the surface feature information must have been consulted in the process of redirecting gaze to the object in the task-relevant spatial location. These results provide further evidence against the claim that the visual system relies solely on spatiotemporal information in establishing object correspondence across saccades, even when it would be advantageous to ignore non-spatiotemporal information.

Although the results support the weak spatiotemporal hypothesis, the results are also potentially consistent with the flexible correspondence hypothesis, which posits that the visual system uses whichever cues are most salient. In Experiments 1 and 2, the target was cued by an event at its location (i.e., the rapid expansion and contraction of the target object), which may have made position information particularly salient in saccade target selection. Experiments 3–5 used non-spatial information to indicate which item was the target so that we could determine whether spatiotemporal information always plays a stronger role than surface feature information, as proposed by the weak spatiotemporal hypothesis, or whether non-spatial information can play a stronger role under conditions that highlight surface features, as proposed by the flexible correspondence hypothesis.

5. Experiment 3

Experiment 3 modified the procedure of Experiment 1 so that the saccade target was cued on the basis of its surface features rather than on the basis of its location. As in Experiment 1, the two array objects appeared to the right of fixation at the beginning of the trial. Rather than cuing one of these objects by means of an event at the location of the object, we cued an object by presenting a color patch at fixation (see Fig. 7). In the surface feature block, participants were instructed to direct gaze to the array object that matched the color of the central patch. In the position block, participants were instructed to direct gaze to the location occupied by the object that matched the central color patch. Thus, whereas the cues in Experiments 1 and 2 directly indicated the location of the target object, even when the color of this object was the relevant feature, the cues in the present experiment indicated the color of the target object, even when the location of this object was the relevant feature.

This manipulation was intended to simulate real-world search situations in which the identity of an object, rather than its location, is known in advance. For example, when searching a cluttered desk for a blue pen, memory for the color and form of the pen can be used to guide attention to objects matching the perceptual features of the pen (e.g., Desimone & Duncan, 1995). In such a case, object surface features would be expected to play a large role in the selection of possible saccade targets. The flexible correspondence hypothesis holds that object correspondence will be sensitive to the informational demands of saccade target selection. If the saccade target is initially selected by virtue of its location, position information will be the
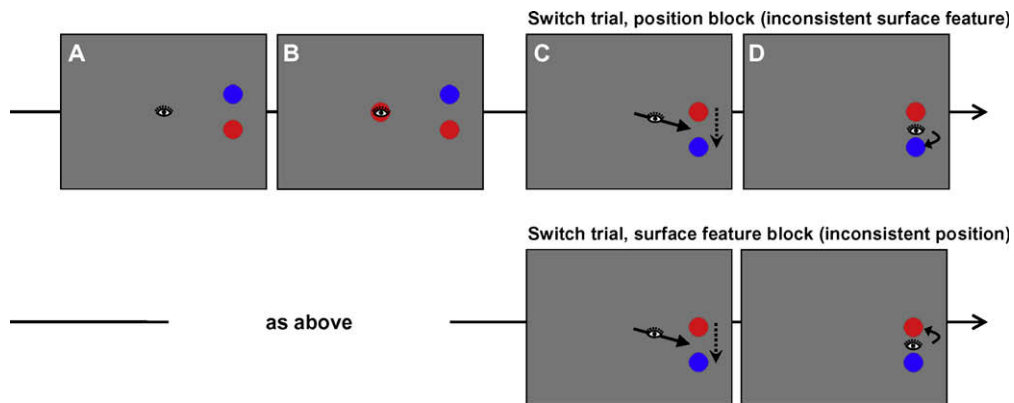
Fig. 7. Sequence of events in the property-switch trials of Experiment 3. Top row: position block. Bottom row: surface feature block. In the grayscale version of this figure, the two colors (red and blue) are represented by white and black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

primary determinant of object correspondence. If the saccade target is initially selected by virtue of its surface features, surface feature information will be the primary determinant of object correspondence. Thus, the flexible correspondence hypothesis predicts surface feature dominance in Experiment 3, because the target is initially specified on the basis of its color. The weak spatiotemporal hypothesis holds that spatiotemporal information always dominates correspondence computations. Thus, the weak spatiotemporal hypothesis predicts position dominance in Experiment 3, similar to that found in Experiments 1 and 2.

5.1. Method

5.1.1. Participants
Sixteen new participants from the University of Iowa community completed the experiment. They either received course credit or were paid. All participants reported normal vision.

5.1.2. Stimuli, apparatus, and procedure
The object stimuli and arrays were the same as in Experiment 1. However, the saccade target was cued by the appearance of a color patch at fixation. Specifically, the two array objects were presented for 1000 ms. Then, a color patch was presented at fixation while the array objects remained visible. Participants executed a saccade to the target as quickly as possible. During the saccade, the central color patch was removed to avoid interference with gaze correction when the eyes landed and to ensure that the correction was made on the basis of memory across the saccade. In all other respects, the procedure was the same as in Experiment 1.

5.1.3. Data analysis
Eyetracking data were analyzed in the same manner as in Experiment 1. A total of 38.7% of the array-shift trials was eliminated for the reasons specified in Experiment 1. For all trials (shift and no-shift), the mean latency of the initial saccade from central fixation to the target object

was 318 ms. These longer initial saccade latencies are consistent with the use of a central, endogenous cue rather than the exogenous cues used in Experiments 1 and 2.

5.2. Results

The results are shown in Fig. 8. Gaze correction on no-switch trials was highly accurate, with 98.5% correct in the position block and 99.3% correct in the surface feature block, $t(15) = 0.78$, $p = .45$. Mean correction latency on no-shift trials was 238 ms in the position block and 219 ms in the surface feature block, $t(15) = 1.49$, $p = .16$.

In the surface feature block, mean correction accuracy on property-switch trials (which introduced inconsistent position information) was reliably lower (91.4%) than accuracy on no-switch trials (99.3%), $t(15) = 3.09$, $p < .01$. And mean correction latency was significantly longer for property-switch trials (289 ms) than for no-switch trials (219 ms), $t(15) = 6.24$, $p < .001$.

In the position block, mean correction accuracy on property-switch trials (which introduced inconsistent surface feature information) was reliably lower (65.9%) than accuracy on no-switch trials (98.5%), $t(15) = 4.62$, $p < .001$. In addition, correction latency was significantly longer for property-switch trials (315 ms) than for no-switch trials (238 ms), $t(15) = 4.28$, $p < .001$. The latency data for the position block must be treated with caution, however, as the low level of accuracy provided few observations for analysis, and the majority of the correct responses were presumably guesses. This high proportion of gaze correction errors was observed despite salient feedback on every trial (the trial did not conclude until the participant had fixated the correct object, at which point a box appeared around that object).

Unlike previous experiments, the interference introduced by inconsistent surface features was larger than the interference generated by inconsistent position. For gaze correction accuracy, the effect of inconsistent position information in the surface feature block (7.9%) was smaller than the effect of inconsistent surface feature information in the position block (32.7%), $t(15) = 2.96$, $p < .01$. For gaze
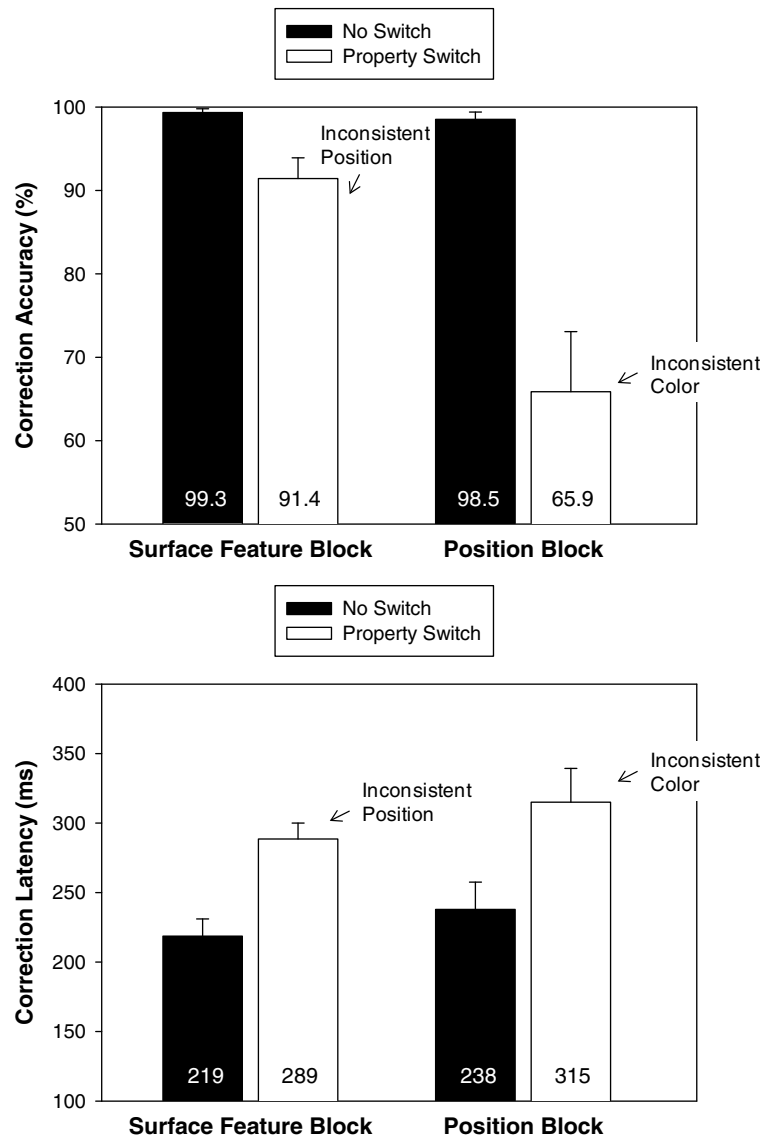
Fig. 8. Experiment 3. Mean gaze correction accuracy (top) and latency (bottom) as a function of the relevant dimension for gaze correction (position block, surface feature block) and property switch. Error bars are standard errors of the means.

correction latency, there was no difference in the magnitude of the consistency effect, with a 77 ms effect of inconsistent surface features and a 70 ms effect of inconsistent position, $t(15) = 0.30$, $p = .77$. Note again that the high proportion of errors in the switch trials of the position block make the latency data difficult to interpret.

5.3. Discussion

In Experiment 3, the saccade target object was initially identified on the basis of its surface features. Contrary to the position dominance observed in Experiments 1 and 2, surface feature information was weighted more heavily than position information in the computation of object correspondence. Thus, the information used to compute object correspondence across saccades varies according to

the demands of saccade target selection, as held by the flexible correspondence hypothesis.

It is important to consider just how strongly these results undermine the hypothesis that spatiotemporal information has primacy in object correspondence across saccades. When correcting gaze on the basis of relative location in the position block, target location before the saccade perfectly predicted target location after the saccade. Participants should have been motivated to constrain the information consulted in gaze correction to the spatial dimension. Nevertheless, switching the two colors led participants to make gaze corrections to the wrong location (which contained the original color) on 34.1% of trials. This is remarkably poor performance for a task that should have been trivial from the perspective of spatiotemporal information.

6. Experiment 4

Experiment 4 replicated Experiment 3 using the real-world object stimuli used in Experiment 2.

6.1. Method

6.1.1. Participants

Sixteen new participants from the University of Iowa community completed the experiment. They either received course credit or were paid. All participants reported normal vision.

6.1.2. Stimuli, apparatus, and procedure

Experiment 4 differed from Experiment 3 in that objects were drawn from the object set of Experiment 2. The saccade was cued by the appearance of one of the two objects at central fixation, as illustrated in Fig. 9.

6.1.3. Data analysis

Eyetracking data were analyzed in the same manner as in Experiment 2. A total of 38.4% of the array-shift trials was eliminated for the reasons specified in Experiment 1. For all trials (shift and no-shift), the mean latency of the initial saccade from central fixation to the target object was 306 ms.

6.2. Results

The results are shown in Fig. 10. Mean gaze correction accuracy on no-switch trials was 96.0% in the position block and 99.1% in the surface feature block, $t(15) = 2.69$, $p < .05$. The source of this difference is not clear, but correction was nonetheless highly accurate in both conditions. Correction latency on no-switch trials was 195 ms in the position block and 188 ms in the surface feature block, $t(15) = 0.92$, $p = .37$.

In the surface feature block, mean correction accuracy on property-switch trials (which introduced inconsistent position information) was reliably lower (78.6%) than accuracy on no-switch trials (99.1%), $t(15) = 6.32$, $p < .001$. And mean correction latency was significantly longer for property-switch trials (281 ms) than for no-switch trials (188 ms), $t(15) = 10.67$, $p < .001$.

In the position block, mean correction accuracy on property-switch trials (which introduced inconsistent surface feature information) was reliably lower (58.2%) than accuracy on no-switch trials (96.0%), $t(15) = 9.48$, $p < .001$. Again, the very high proportion of gaze correction errors was observed despite the fact that participants received feedback on every trial. Mean correction latency was significantly longer for property-switch trials (256 ms) than for no-switch trials (195 ms), $t(15) = 3.50$, $p < .005$.

For gaze correction accuracy, the size of the effect of inconsistent position information (20.5%) was smaller than the size of the effect of inconsistent surface feature information (37.8%), $t(15) = 3.55$, $p < .005$, replicating Experiment 3. For gaze correction latency, there was no difference in the magnitude of the consistency effect, with a 61 ms effect of inconsistent surface features and a 93 ms effect of inconsistent position, $t(15) = 1.63$, $p = .12$. As in Experiment 3, the large differences in accuracy, and the fact that 41.8% of switch trials were eliminated from the surface feature block due to inaccurate correction, limit interpretation of the latency data.

6.3. Discussion

In Experiment 4, participants selected the saccade target on the basis of its surface feature match with a centrally presented real-world object. Gaze correction was significantly more impaired by inconsistent surface feature information than by inconsistent position, replicating Experiment 3. The effect of surface feature consistency in the position block was observed despite the fact that the participants should have been motivated to respond solely on the basis of spatio-temporal information. Yet, fully 41.8% of corrections were directed to the wrong location (which contained the original real-world object). These data provide further evidence that the information functional in establishing object correspondence across saccades is determined flexibly by the task, consistent with the flexible correspondence hypothesis.

7. Experiment 5

Experiment 3 and 4 provided evidence that object correspondence across saccades can weight surface feature information more heavily than position information. These
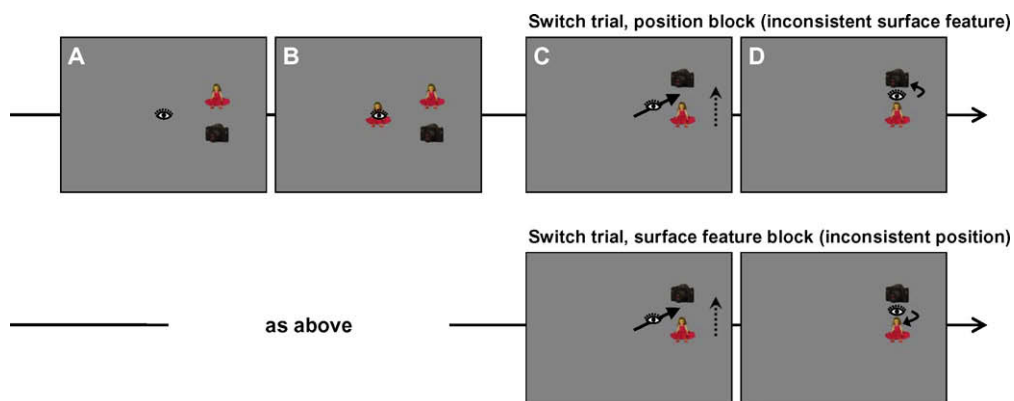

Fig. 9. Sequence of events in the property-switch trials of Experiment 4. Top row: position block. Bottom row: surface feature block.
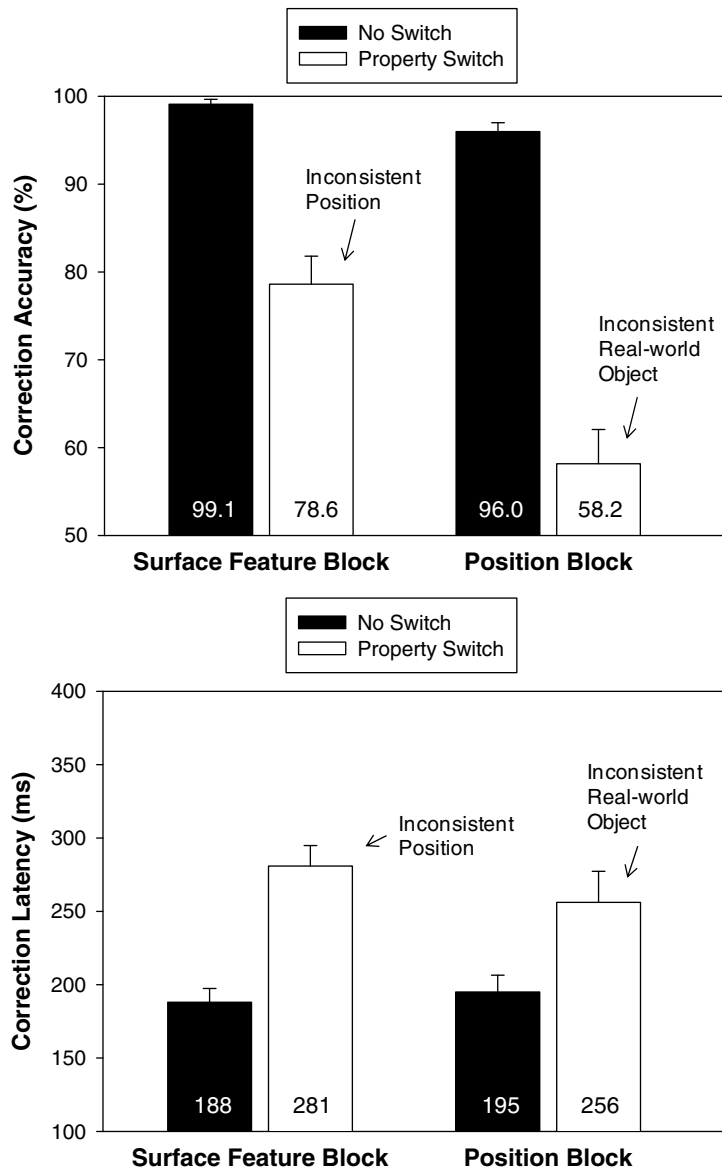
Fig. 10. Experiment 4. Mean gaze correction accuracy (top) and latency (bottom) as a function of the relevant dimension for gaze correction (position block, surface feature block) and property switch. Error bars are standard errors of the means.

results reversed the effects seen in Experiments 1 and 2, in which position information was weighted more heavily. The reversal of the effect could be generated by two related causes. First, the mere presence of a spatially local transient signal in Experiments 1 and 2 could have highlighted position information in the task. Second, and as we have argued thus far, the critical difference could be the nature of the information used to select the saccade target. In this latter view, any task in which the target is initially selected on the basis of location should generate position dominance.

To tease apart these possibilities, Experiment 5 employed an arrow cue presented at central fixation (see Fig. 11). Participants saw two color disks (as in Experiments 1 and 3). Then, an upward- or downward-pointing arrow was presented at fixation. The participants' task was to fixate the object in the position cued by the arrow (position block) or the object with the color appearing at the cued location (surface feature block). The central arrow was an endogenous cue that did not create a transient signal at the cued location, yet the target object was specified on the basis of its position. If object correspondence and gaze correction are driven by the nature of the information used to select the saccade target, then position dominance should be observed despite the absence of a transient signal at the target location.
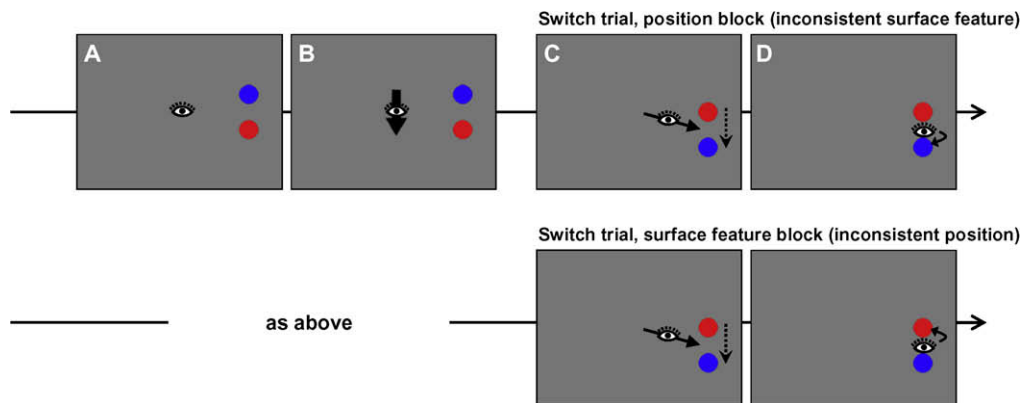
Fig. 11. Sequence of events in the property-switch trials of Experiment 5. Top row: position block. Bottom row: surface feature block. In the grayscale version of this figure, the two colors (red and blue) are represented by white and black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

7.1. Method

7.1.1. Participants

Sixteen new participants from the University of Iowa community completed the experiment. They either received course credit or were paid. All participants reported normal vision.

7.1.2. Stimuli, apparatus, and procedure

Experiment 5 cued the target object by an upward- or downward-pointing arrow presented at fixation (see Fig. 11). The arrow subtended 1.6° of visual angle. During the saccade to the object array, the arrow was removed.

7.1.3. Data analysis

Eyetracking data were analyzed in the same manner as in Experiment 1. A total of 38.6% of the array-shift trials was eliminated for the reasons specified in Experiment 1. For all trials (shift and no-shift), the mean latency of the initial saccade from central fixation to the target object was 304 ms. Thus, the endogenous arrow cue generated initial saccade latencies similar to those observed for the endogenous color patch and natural object cues used in Experiments 3 and 4.

7.2. Results

The results are shown in Fig. 12. Mean gaze correction accuracy on no-switch trials was 100% in the position block and 95.2% in the surface feature block, $t(15) = 1.77$, $p = .10$. Mean correction latency on no-switch trials was 224 ms in the position block and 234 ms in the surface feature block, $t(15) = 1.26$, $p = .23$.

In the surface feature block, mean correction accuracy on property-switch trials (which introduced inconsistent position information) was reliably lower (70.7%) than accuracy on no-switch trials (95.2%), $t(15) = 4.94$, $p < .001$. And mean correction latency was significantly longer for property-switch trials (331 ms) than for no-switch trials (234 ms), $t(15) = 9.75$, $p < .001$.

In the position block, mean correction accuracy on property-switch trials (which introduced inconsistent sur-

face feature information) was equivalent with correction accuracy on no-switch trials (both 100%). In addition, there was no effect of surface feature consistency on correction latency, with mean correction latency of 225 ms for property-switch trials and 224 ms for no-switch trials, $t(15) = .26$, $p = .79$.

The effect of position consistency in the surface feature block was larger than the effect of surface feature consistency in the position block, both for correction accuracy, $t(15) = 4.94$, $p < .001$, and latency, $t(15) = 7.73$, $p < .001$.

7.3. Discussion

In Experiment 5, a central arrow was used to cue the saccade target position without a transient signal at the target location. Position information dominated object correspondence and gaze correction operations. Thus, a transient signal at the target location is not necessary to generate position dominance in object correspondence across saccades. In general, if the saccade target object is selected by virtue of its position, position is preferentially weighted in correspondence. However, if the target is selected by virtue of its surface features, surface features are preferentially weighted (Experiments 3 and 4). Note that when the target was cued by an arrow, the effect of surface feature consistency was eliminated entirely. Thus, it appears that participants can limit transsaccadic memory and/or comparison operations to position information when the task is configured so that the target is selected solely on the basis of position. In contrast, the significant spatial interference generated in Experiments 3 and 4 suggests that participants cannot limit transsaccadic memory to surface feature information. This asymmetry likely reflects the fact that position encoding is mandatory in motor programming; one cannot program a saccade to an object without encoding its location.

8. General discussion

Visual perception is continually disrupted by saccades, blinks, and occlusion. The present study examined how visual memory is used to establish the correspondence
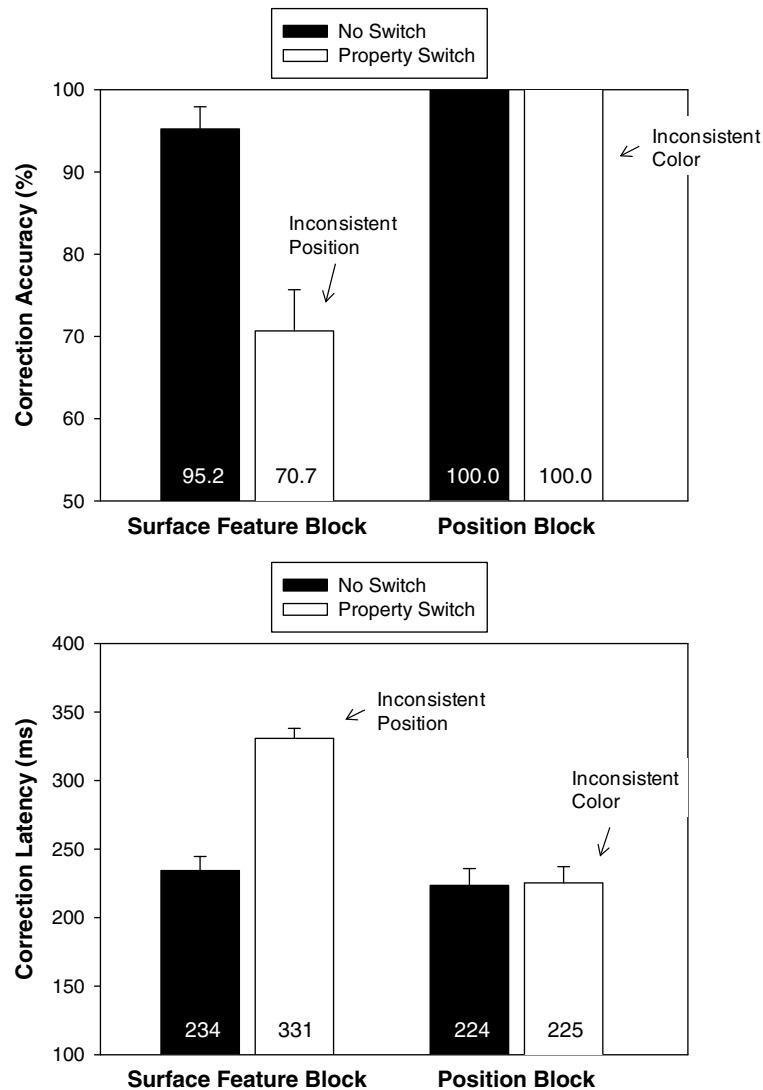
Fig. 12. Experiment 5. Mean gaze correction accuracy (top) and latency (bottom) as a function of the relevant dimension for gaze correction (position block, surface feature block) and property switch. Error bars are standard errors of the means.

between objects visible before and after a disruption. We chose to examine object correspondence across eye movements. Saccades are by far the most frequent form of perceptual disruption, and saccades pose significant challenges to establishing object correspondence, because the retinal locations of objects change across a saccade. Theoretical approaches to object correspondence have been dominated by the object-file theory of Kahneman et al. (1992), which holds that objects are addressed by their spatial locations. Memory for other properties of an object (e.g., surface features, identity) may be bound to a spatial index marking the remembered object location, but the location information is primary. In this view, object correspondence across brief disruptions, such as saccades, depends entirely on spatial continuity and does not consult other possible identifying information (such as surface features or object identity). Yet, the Kahneman et al. claim of

position dominance in object correspondence operations has not been tested extensively. In the present study, we systematically examined the roles of position and surface feature information in computing object correspondence across saccades.

To this end, we developed a paradigm that simulated the real-world situation in which a saccadic eye movement misses a target object, leading to a memory-guided corrective saccade, perhaps the most common real-world situation in which visual memory is used to compute object correspondence (Hollingworth et al., 2008). Gaze errors were created experimentally by shifting two array objects during a saccade to one of them. This caused the eyes to land between the objects, and a corrective saccade was required to bring the eyes to the target. On a subset of these trials, the two objects switched properties during the saccade so that we could dissociate location and surface

feature information, both of which could potentially guide gaze corrections in the natural environment. We tested a strong spatiotemporal hypothesis (that correspondence is established solely on the basis on spatiotemporal information), a weak spatiotemporal hypothesis (that surface features are also consulted, but spatiotemporal information is primary), and a flexible correspondence hypothesis (that the two sources of information are weighted flexibly on the basis of task demands).

When the task required participants to direct gaze to an object defined by its surface features, gaze correction accuracy and efficiency were impaired when the relative position of the target changed during the saccade. This finding is equivalent to the consistent-position benefit observed in previous object-file studies (Kahneman et al., 1992), and it is consistent with the hypothesis that correspondence is determined solely on the basis of spatiotemporal information. However, when the task required participants to direct gaze to an object defined by its location, gaze correction accuracy and efficiency were impaired when the surface features of the target object changed during the saccade. This result demonstrates that surface feature information is consulted in the computation of correspondence across saccades. Therefore, the results falsify the strong spatiotemporal hypothesis (Kahneman et al., 1992), at least within the domain of correspondence across saccades.

In Experiment 5, the saccade target was specified solely on the basis of its spatial position, and spatial consistency dominated object correspondence. In Experiments 3 and 4, the saccade target was specified on the basis of its surface features, and surface-feature consistency dominated object correspondence. The reversal of the dominant information used to establish object correspondence can be accommodated only by the flexible correspondence hypothesis. Specifically, the information used to establish object correspondence across saccades is determined flexibly, with the relative weighting of information governed by the demands of saccade target selection.

For example, if one searches for a favorite blue pen, and a particular peripheral object is selected as the saccade target because it matches well the surface features of the pen (blue, cylindrical, elongated, relatively small), then these surface feature properties are likely to be preferentially encoded into VSTM and stored across the saccade.[2] When the eyes land, the mapping of objects visible before and after the saccade will be determined primarily by surface feature consistency. In particular, the continuity between the pen perceived in the periphery and the pen now falling near the fovea will be established on the basis of a surface feature match between the remembered properties of the pen and the properties of the pen perceived after the saccade. If a corrective saccade is required, it will be directed to the object matching the remembered surface feature properties (Hollingworth et al., 2008).

In contrast, if one's attention is drawn to some unknown object skittering beneath a nearby table in a restaurant, then the saccade target is likely to be selected on the basis of its salient spatiotemporal properties (that it is moving beneath the table), and these properties will be preferentially encoded into VSTM and stored across the saccade. When the eyes land, correspondence will be established on the basis of a match between the spatiotemporal properties retained in memory and the spatiotemporal properties of objects visible after the saccade. If a corrective saccade is required, the saccade will be directed to the object that best matches the remembered spatial properties (i.e., to an object that is under the table).

In more general terms, we propose that before a saccade, the saccade target is attended (e.g., Hoffman & Subramaniam, 1995), and perceptual properties of the target that drove its selection are preferentially activated and consolidated into VSTM. These properties are then maintained in VSTM across the saccade (Irwin, 1992). When they eyes land, objects falling near the saccade landing position are compared with the contents of VSTM (Currie et al., 2000; Hollingworth et al., 2008). The nature of the information entering into this comparison operation – i.e., the information used to establish correspondence – will necessarily depend on the type of information stored across the saccade in VSTM. If surface features are preferentially encoded and retained in VSTM across the saccade, then surface feature information will play the central role in comparison after the saccade. If spatiotemporal information is preferentially encoded and retained in VSTM across the saccade, then spatiotemporal information will play the central role in comparison.

Note that the above account assumes that differential weighting of surface feature and spatiotemporal information occurs at the stage of encoding object properties into VSTM, with the object properties relevant for saccade target selection preferentially encoded and retained across the saccade. An alternative is that differential weighting occurs not at encoding but at the comparison stage, with certain object properties prioritized in the comparison of VSTM representations to perceptual information after the saccade. The present data do not distinguish between these two alternatives, and they need not be mutually exclusive.

In addition, we are agnostic with respect to whether surface feature and spatiotemporal correspondence are supported by a single VSTM system or by two specialized systems. There is significant evidence suggesting that VSTM is composed of partially independent subsystems for the representation of object perceptual form and object location (for a review, see Luck, in press). This raises the clear possibility that there are multiple memory systems capable of computing object correspondence: (1) an object VSTM system that computes correspondence on the basis of ventral stream features such as color, shape, and identity and (2) a spatial VSTM system that computes correspondence on the basis of dorsal stream features such as position and motion.[3] If this is indeed the case, then the

---

[2] In fact, it is likely that remembered surface features of a search target would be maintained in VSTM throughout search so that perceptual features of each attended object could be compared with remembered properties of the target (Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Woodman, Luck, & Schall, 2007).

[3] Others researchers use the terms "visual working memory" and "spatial working memory" to refer to these systems.

present results would demonstrate that both dorsal and ventral systems support object correspondence across saccades and that the relative contributions of the two systems are weighted according to the demands of saccade target selection.

8.1. Relationship between the present study and previous work on object correspondence

If surface feature information can be used to establish object correspondence across saccades, why have studies using the standard object-reviewing paradigm found no effect of surface feature consistency (Kahneman et al., 1992; Mitroff & Alvarez, 2007)? The Mitroff and Alvarez study provided the strongest test of surface feature effects in the standard object-reviewing paradigm. In their experiments, participants first saw two objects. In the spatiotemporal condition, the objects were identical, simple boxes. In the surface feature condition, the objects were more complex and differed on a number of surface feature attributes. Letters appeared briefly in the two objects. In the spatiotemporal condition, the two "empty" objects moved smoothly to new positions. In the surface feature condition, the two "empty" objects were moved in a single step to the new positions (making spatial information ambiguous). A test letter was 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. When the test letter matched one of the preview letters, the test letter could appear either in the correct or incorrect position (spatiotemporal condition) or within the object with the correct or incorrect surface features (surface feature condition). A consistency effect was observed only in the spatiotemporal condition.

One limitation of this design, however, is that the two conditions were not equated for discontinuity on the two dimensions. In the spatiotemporal condition, there was no surface-feature discontinuity; the two objects kept 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. Objects in the world do not appear in a new position without traversing a path from the original location to the new location. Thus, the salient spatiotemporal discontinuity in the surface feature condition could have masked an effect of surface feature consistency. To provide an appropriate comparison between the two conditions, one would need to either (1) eliminate spatial discontinuity in the surface feature condition or (2) include a surface feature discontinuity in the spatiotemporal condition (e.g., by changing the surface features of the two objects during their motion).

The design of the present experiments avoided these problems. When objects switched properties, the perceptual discontinuity on each dimension was the same in the surface feature block and the position block, because the physical stimuli were identical in the two blocks (see Fig. 1); the only difference was the task participants performed. In addition, the transient signal generated by the switch itself was masked, because it occurred during the saccade, reducing the salience of visual discontinuity on

both dimensions. That is, in the surface feature condition of Mitroff and Alvarez (2007), the position change happened during a fixation and was directly visible, potentially generating salient evidence against object continuity. In the present study, however, the property switch occurred during a saccade, when it was not directly visible, allowing us to manipulate surface feature and spatial consistency without introducing salient, visible discontinuity on either dimension.

There are two additional differences between our study and previous studies that could potentially account for the discrepant results. First, our experiments had no visible motion in any condition, whereas the traditional object-reviewing paradigms have typically depicted objects in motion. Perhaps surface feature effects are limited to situations depicting non-moving objects. Second, our experiments examined correspondence across saccades, whereas Kahneman et al. (1992) and Mitroff and Alvarez (2007) examined correspondence operations across motion and spatial displacement. Perhaps surface features are consulted only in transsaccadic correspondence operations. More generally, our experiments did not probe whether participants consciously perceived object persistence as a function of surface feature and spatial consistency. Perhaps surface features can be used to map objects across disruptions, but surface feature consistency might not influence one's conscious perception of object persistence.

Two recent studies provide initial evidence against all three of these accounts. First, Moore, Mordkoff, and Enns (2007; Moore & Enns, 2004) presented participants with a dynamic object display in which an object disk was shifted in a series of discrete steps around a virtual circle. Participants typically perceived a single object moving in a circular path. However, if the color or the size of the object was changed in one of the frames, participants perceived two objects: a trace of the original object at the previous location and a new, different object at the current location. Thus, when color and size consistency were maintained, an object at a new location was considered a continuation of the object perceived earlier, and only one object was perceived. However, if there was a salient change in surface features, the new object was not considered to be a continuation of the object perceived earlier, and participants saw two objects instead of one. These data complement the present results within a different perceptual domain requiring correspondence. Specifically, the Moore et al. (2007; Moore & Enns, 2004) results demonstrate that (1) surface feature effects are observed in paradigm that depicts objects in motion, (2) the use of surface features in object correspondence is not limited to correspondence across saccades, and (3) surface feature consistency directly influences conscious perception of object persistence. Although Moore, Mordkoff, and Enns (2007) showed that under some circumstances, surface feature information can trump spatiotemporal information in an apparent motion paradigm, it is important to note that the more common finding in the apparent-motion literature is greater weighting of spatiotemporal information than surface feature information (for a review, see Dawson,

1991). A precise account of the relative contributions of spatiotemporal and surface feature information in this domain must await further research.

Second, Hollingworth and Franconeri (in preparation) modified the Mitroff and Alvarez (2007) method in a manner that eliminated visible discontinuity on each dimension. Two color 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 at different vertical positions. 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 different shape. When both shapes were the same as in the preview, spatial and color consistency were manipulated independently. (Because the object manipulations were introduced during occlusion, there was no visible discontinuity in either color or position.) The results replicated those in the present study. Both color and position consistency effects were observed, and they were approximately additive. Thus, surface feature effects are observed even in a close analog of the original object reviewing paradigm. Sensitivity to surface feature consistency appears to be a general property of object correspondence mechanisms.

8.2. A special role for spatiotemporal information in occulomotor behavior?

The preceding discussion has stressed that object correspondence operations use multiple sources of perceptual information and are not limited to spatiotemporal information. However, there is at least one way in which spatial information plays a special role in object correspondence. In the present gaze correction task, when the saccade target was selected by virtue of its surface features (Experiments 3 and 4), there remained significant interference from inconsistent position, suggesting that position was coded into VSTM across the saccade. However, when the saccade target was selected by virtue of its position (Experiment 5), there was no interference from inconsistent color, suggesting that participants were able to ignore color information. This asymmetry is likely to be caused by the inherently spatial nature of the saccade task. The results from Experiment 5 suggest that it is possible to generate a saccade to an object location without encoding the surface features of the object at that location. However, it is never possible to execute a saccade to an object without encoding its location, because the occulomotor component of the task requires that a saccade be generated to a particular position, accounting for the significant position interference observed in Experiments 3 and 4. Thus, in any task involving motor interaction with an object, position encoding is mandatory, but surface feature encoding need not be mandatory.

8.3. Implications for the role of position in visual object cognition

What implications do the present results have for understanding the role of position in object correspondence and, more generally, in object perception and memory?

The original framework developed by Treisman and colleagues placed spatial position at the very center of visual cognition. Attending to location was proposed to be required for the binding of different surface features of an object into a coherent perceptual representation (Treisman & Gelade, 1980). Attention to object location was also proposed to be necessary for the maintenance of feature binding in VSTM (Wheeler & Treisman, 2002). Moreover, gaining access to the remembered features of an object was proposed to require attending to the location to which memory for object properties was bound (Kahneman et al., 1992). Finally, assigning spatial indexes to objects was proposed to play an important role in specifying particular tokens of a conceptual type (e.g., distinguishing *that* dog from among other visible dogs).

Although spatial position is indeed an efficient form of representation for property binding and reference, subsequent research has shown that solutions to problems of binding and reference in visual memory need not be limited to spatial attention and spatial indexing. First, attention to object location is not necessary to maintain feature binding in VSTM. In tasks requiring memory for the binding of perceptual features (such as a particular combination of color and orientation), the withdrawal of attention from the locations of items in VSTM does not specifically impair binding memory compared with memory for individual features (Gajewski & Brockmole, 2006; Johnson et al., 2008); coherent object representations are preserved after attention has been directed elsewhere (Hollingworth, 2004; Hollingworth & Henderson, 2002). Thus, although attention to location is likely to play an important role in initial perceptual binding, once bound, object representations can remain bound in VSTM without continued focal attention to object location.

Second, memory for features in VSTM can be accessed to some extent independently of object positions. If objects were addressed only by their positions, and surface feature information could be accessed only by activating the location to which those features were bound (Kahneman et al., 1992), changes in the positions of objects should place severe limitations on the retrieval of object surface features. Although changing the positions of objects in VSTM tasks interferes to some extent with the retrieval of featural information (Hollingworth, 2007; Jiang, Olson, & Chun, 2000; Treisman & Zhang, 2006), such effects are only moderate, and participants can reliably access memory for features despite changes in object position. These data suggest that visual memory representations are not addressed solely by their locations; they can be addressed by surface-feature content as well. If object memory in VSTM can be addressed by surface feature content, then surface features can be used to establish object correspondence across perceptual disruptions, as found in the present experiments.

Third, the claim that position indexes are necessary to individuate tokens of a particular conceptual type is likely to have a more limited application than that originally intended by Kahneman et al. (1992). When object-file theory was developed, theoretical approaches to object memory and recognition held that object representations were

highly abstract and would not be sufficient to distinguish individual tokens of a particular category (e.g., Biederman, 1987). Thus, similar objects (e.g., two different dogs) would lead to the same object description, and position indexes would be necessary to individuate them. Similarly, the phenomenon of change blindness has led some researchers to conclude that very little perceptual detail is remembered from objects, and spatial indexes are therefore necessary to refer to them (Pylyshyn, 2000; Spivey, Richardson, & Fitneva, 2004). However, subsequent work has shown that object representations in memory can preserve considerable visual detail (for reviews, see Hollingworth, 2008; Palmeri & Tarr, 2008). For example, in Hollingworth (2004, 2005), participants could remember the token version of a particular object in a scene (e.g., choose correctly which of two kettles had appeared in a kitchen) both during online scene viewing and after delays as long as 24 h. If visual memory for surface features is sufficient to discriminate object tokens, and if object memory can be addressed by surface feature content, then one does not necessarily need a spatial index to make reference to an object token (except in the fairly rare case that there a multiple, nearly identical objects in the scene; Pylyshyn & Storm, 1988).

Finally, the present data raise the question of whether there remains a meaningful distinction between object correspondence and object recognition. Both object recognition and object correspondence describe situations in which a current sensory input is linked with a previous sensory input, and both may involve using remembered spatiotemporal information and surface feature information to make such links. In the original Kahneman et al. (1992; see also Treisman, 1992) account, the proposed difference between object correspondence and object recognition was that object recognition depends on LTM for location-independent object form, whereas object correspondence is computed over very short delays, depends on spatiotemporal continuity rather than surface features, and generates the conscious experience of object persistence. The present data demonstrate that over extremely short disruptions (saccades that last between 30 and 50 ms), object mapping is informed not just by spatiotemporal continuity but also by surface feature continuity. In addition, the Moore et al. (2007) study demonstrates that one's conscious perception of object persistence is similarly informed both by spatial information and by surface feature information. Across longer timescales typically associated with object recognition, the identification of individual objects certainly depends on memory for surface feature information (Palmeri & Tarr, 2008). However, objects are not stored independently of locations in LTM, and memory for position contributes to recognition: recognition memory for objects is impaired if the object is presented at a different scene location than the one in which it appeared at study (Hollingworth, 2006, 2007), which is an LTM analog of the position consistency effect found by Kahneman et al. (1992) over brief delays. Thus, although the subjective experience of continuity may occur only across brief delays, there does not appear to be an exclusive division in the use of position and surface feature information as a function of the timescale over which comparison occurs or as a function of conscious perception. As a result, current evidence does not compel a strong distinction between object correspondence and object recognition.

References

Becker, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Vision and visual dysfunction. Eye movements* (Vol. 8, pp. 93–137). London: MacMillan.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.

Currie, C., McConkie, G., Carlson-Radvansky, L. A., & Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. *Perception & 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.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Deubel, H., Wolf, W., & Hauske, G. (1982). Corrective saccades: Effect of shifting the saccade goal. *Vision Research, 22*, 353–364.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition, 97*(3), 295–313.

Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review, 13*, 581–587.

Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General, 123*, 410–426.

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 826–839.

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.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*, 787–795.

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. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 396–411.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 58–69.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 31–47.

Hollingworth, A. (2008). Visual memory for natural scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory*. New York: Oxford University Press.

Hollingworth, A., & Franconeri, S. L. (in preparation). The role of surface features in establishing object correspondence across motion and occlusion.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 113–136.

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 & Psychophysics, 69*, 172–184.

Hyun, J.-S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (in press). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance*.

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. (1996). 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.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 683–702.

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, 34*, 41–55.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46*, 1762–1776.

Luck, S. J. (2008). Visual short-term memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory*. New York: Oxford University Press.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*, 899–917.

Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, underlie object persistence. *Psychonomic Bulletin & Review, 14*, 1199–1204.

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.

Palmeri, T. J., & Tarr, M. J. (2008). Visual object perception and long-term memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 163–208). New York: Oxford University Press.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Pylyshyn, Z. (2000). Situating vision in the world. *Trends in Cognitive Science, 4*, 197–207.

Pylyshyn, Z. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition, 11*(7), 801–822.

Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 1–19.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*, 259–290.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*, 159–177.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. (1999). When are spatiotemporal and featural properties encoded as a result of attentional allocation? *Association for Research in Vision and Ophthalmology*, 5/13/99, Ft. Lauderdale, FL.

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.

Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychologist, 47*, 862–875.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A., & Zhang, W. W. (2006). Location and binding in visual working memory. *Memory & Cognition, 34*, 1704–1719.

Van de Walle, G., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development, 1*, 249–280.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*, 48–64.

Woodman, G. F., Luck, S. J., & Schall, J. D. (2007). The role of working memory representations in the control of attention. *Cerebral Cortex, 17*, 118–124.

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 & Psychophysics, 67*, 676–690.