Cognition 276 (2026) 106657

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Full Length Article

The tiger moving through the tall grass: Spontaneous perception and tracking of change-defined objects from second-order features

Dawei Bai [a,*], Brian J. Scholl [a,b,*]

[a] Department of Psychology, Yale University, United States of America
[b] Wu Tsai Institute, Yale University, United States of America

ARTICLE INFO

Keywords:
Object perception
Object tracking
Multiple-object tracking
Binding problem
Second-order motion

ABSTRACT

The most important objects — from predators to projectiles — are often those that are moving. Accordingly, visual systems are specialized for object tracking, and many models of such processing involve continually re-identifying surface features across space and time (as you might follow a tiger by keeping track of its orange stripes). Here, in contrast, we show across four experiments — and in phenomenologically compelling demonstrations — that people also spontaneously perceive and track *change-defined* objects, with no enduring surface properties from moment to moment. Observers viewed regular grids filled with hundreds of small elements (e.g. randomly oriented crosses). Change-defined objects were implemented by having an element change (e.g., from one random orientation to another random orientation), with these changes propagating through space and time. Observers still detected the 'second-order' motion in such displays — despite the lack of persisting features, and without being able to identify the object in any static frame. And beyond motion detection, observers also spontaneously perceived and tracked persisting objects — and were even able to perform multiple-object tracking. We also generalized this phenomenon in several ways, showing that it occurs with other types of changes (e.g., to brightness or shape), and even when different types of changes are haphazardly interleaved (e.g. with a random orientation change to one element followed by a random brightness change to its neighbor, etc.). In essence, these stimuli show how observers can track a tiger that is always fully hidden while moving through tall grass, by tracking variable local changes to the grass itself.

1. Introduction

Visual perception is critical for mental life not just in characterizing static scenes, but in predicting what is about to happen. Accordingly, vision must be able not only to segment discrete objects from the background, but also to track them over time and space — as when viewing a moving predator, or a thrown projectile — and extensive research has confirmed that visual systems are specialized in various ways for tracking dynamic objects (Holcombe, 2023; Scholl, 2009). This foundational ability emerges remarkably early in development (Bower et al., 1971; Bremner et al., 2005; Johnson et al., 2003), and in adults it operates reliably even in challenging situations — as when tracking (a) flying balls at high speeds (Bahill & LaRitz, 1984; McLeod, 1987), (b) objects moving unpredictably (Tavassoli & Ringach, 2010), and (c) even *multiple* randomly moving objects among distractors (Holcombe, 2023; Pylyshyn & Storm, 1988; Scholl, 2009). These sophisticated abilities seem crucial for perception and navigation in dynamic environments. As such, object tracking has also become a central focus in AI work — e.g. in applications such as self-driving cars and surveillance systems (e.g. Luo et al., 2021; Milan et al., 2016; Zhang et al., 2008).

Models of object tracking — from infant cognition, to adult psychophysics, to computer vision — often share a common theme: such processing starts by first identifying some surface features of an object, and the *persistence* of these features then allows the visual system to track it in space and time. The tiger in Fig. 1, for example, might be tracked by identifying and then continuously *re*identifying its orange color, or its characteristic stripes. (And even when tracking an object amidst identical distractors — say, when watching a particular ant moving among seemingly identical ants — one may still reidentify the same ant-like features in a more constrained set of locations from moment to moment.)

* Corresponding author at: Department of Psychology, Yale University, Box 208047, New Haven, CT 06520-8047, United States of America
E-mail addresses: dawei.bai@hotmail.com (D. Bai), brian.scholl@yale.edu (B.J. Scholl).

https://doi.org/10.1016/j.cognition.2026.106657
Received 4 February 2026; Received in revised form 7 July 2026; Accepted 9 July 2026
Available online 31 July 2026

1.1. Introducing change-defined objects

The present study, in contrast, explored the ability to track entities that may at first blush seem somehow impossible: objects which have no enduring surface properties from one moment to the next, and so are defined only by pure change itself. To create such objects, we adapted stimuli from a form of so-called 'second-order motion', often defined by a type of stimulus flicker (Badcock & Derrington, 1985; Chubb & Sperling, 1988). Our change-defined object displays consist of hundreds of small elements arranged in a regular grid — e.g. randomly oriented crosses (Fig. 2). A moving object is implemented by having one element change from its original random orientation to a new random orientation — after which a *different* change (i.e. to another random orientation) occurs to a neighboring element, and so forth — with the locations of these random changes propagating through space and time. Such change-defined objects can also be implemented via random changes to other properties, such as brightness or shape (Fig. 3).

Change-defined objects thus differ from typical objects in two related ways. First, they cannot be tracked via enduring surface features, since such features do not exist. (Of course each cross on each frame still has particular features such as orientation — and the changes themselves could be considered as types of features [Mazalik et al., 2023], but none of these persist as stable surface features.) You can follow the tiger in Fig. 1 by (re)identifying its specific color, but you cannot track the change-defined object in Fig. 2 by (re)identifying its specific orientation — because that orientation is in constant flux. And because that flux is random, one cannot even track a change-defined object by noting a consistent *amount* of change from cross to cross (e.g. when tracking a rotating object by noting that its orientation is changing by a regular predictable degree from moment to moment). Second, change-defined objects cannot even be segmented from the background texture in the first place based on surface features: in a single static frame, it is impossible *even in principle* to locate a change-defined object, since it is surrounded by similar-looking elements. (A moving tiger is easy to depict, as in Fig. 1 — where its unique color distinguishes it from the background. Even just *depicting* change-defined objects in static figures is thus challenging, requiring some work — as in Fig. 2.)

In sum, change-defined objects could be tracked only by the spatiotemporal continuity of transient changes themselves. Essentially, such stimuli require observers to perceive and track a tiger that is always fully hidden while moving through tall grass, by tracking variable local changes to the grass itself over time and space.

1.2. Relation to previous work: Object tracking, texture segregation, apparent motion, and second-order motion

Change-defined objects seem distinct from related cases of object tracking along theoretically important dimensions. Perhaps the most relevant case is the classic 'tunnel effect' (e.g. Burke, 1952; Flombaum et al., 2004): when a red disc moves on a spatiotemporally smooth path behind a screen, re-emerging (at the right place and time) as a green square, people have no trouble tracking it (Flombaum & Scholl, 2006). This demonstrates that *occasional* featural changes do not disrupt object tracking during periods (typically much longer) *without* occlusion — when an object maintains its properties from moment to moment. Change-defined objects, in contrast, shift their appearance both instantaneously and continuously — such that trajectories must be extrapolated *only* from changes.

The current stimuli are also reminiscent of the prominent literature on texture segregation, in a few related ways. First, when looking at a random-noise field whose quadrants differ in statistical properties (e.g. first- or second-order statistics), viewers can readily segment the quadrants, even though the value in each pixel is random across the entire field (Julesz, 1962; Julesz & Bergen, 1983). Relatedly, in a field of randomly oriented Gabor patches, viewers can easily pick out the set of aligned patches that outline contours of real-world objects (e.g. a sock; Nygård et al., 2009, 2011). However, such texture segregation and contour detection do not rely on motion, since the different quadrants and objects are distinguishable even from a static frame — whereas a change-defined object is visible only in dynamic displays. Second, another related sort of phenomenon does involve motion in the form of 'motion-defined objects' — as when random-noise patches move across a random-noise background (Regan & Beverley, 1984; Regan & Hamstra, 1991; Reichardt & Poggio, 1979). While in motion, such objects can be readily tracked just as any other object (Flombaum et al., 2008; Gao & Scholl, 2010) — even though they are also invisible in static frames. Critically, however, motion-defined objects always maintain their surface features (even if those features are complex): it is always the *same* random-noise patch that is moving. (In this sense, tracking a motion-defined object is like tracking a tiger in a savannah that is densely populated with many other tigers.) This is also relevant to the sorts of non-rigid 'vortex' motions that have been studied in recent work on eye movements: these displays involve random-dot fields in which the dots rotate cohesively inside a moving circular region, with the resulting moving vortices seeming smooth when viewed peripherally, even though observers cannot track the vortices using smooth pursuit eye-movements (Koerfer et al., 2024; Koerfer & Lappe, 2022). Here the vortices are also invisible in static frames, but the intra-vortex motion is itself predictable (e.g. rotating in the same direction) from moment to moment. Change-defined objects, however, have nothing constant from one frame to the next.

Perceiving change-defined objects in displays such as Fig. 2 clearly represents a form of apparent motion. Indeed, this seems true by definition, given the spatial separation of the stimuli. But apparent motion by itself does not explain or entail the perception of change-defined objects, since previous work in this domain has never tested these types of stimuli. Even when two frames in an apparent motion display are different, they are always readily segmentable from the background (e.g. Cavanagh & Mather, 1989; Mather & Anstis, 1995).

Perhaps most important, our stimuli are clearly a variety of 'second-order' motion (Badcock & Derrington, 1985; for a review, see Nishida, 2011): for example, when viewing translating bars composed of noise dots, with each dot randomly presented as black or white on each frame, observers can still discriminate the global motion direction on a static

Fig. 1. A tiger (not in any tall grass!) might be readily tracked simply by identifying and then continuously reidentifying its orange color on the mostly-green background.

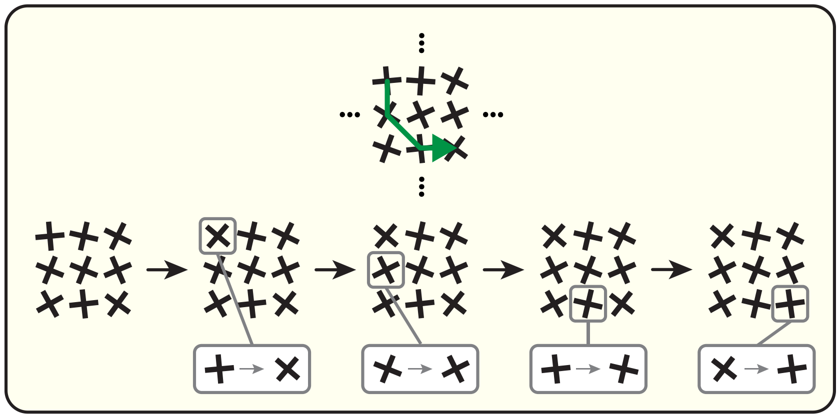

Fig. 2. A dynamic change-defined object, implemented as changes to a cross orientation (top), with a frame-by-frame depiction of its motion (bottom). On each frame, one cross changes its random orientation to a new random orientation, and such changes propagate from one cross to another (with every change resulting in a new random orientation), while maintaining spatiotemporal continuity. This gives rise to the percept of a single object in motion, defined by pure change.

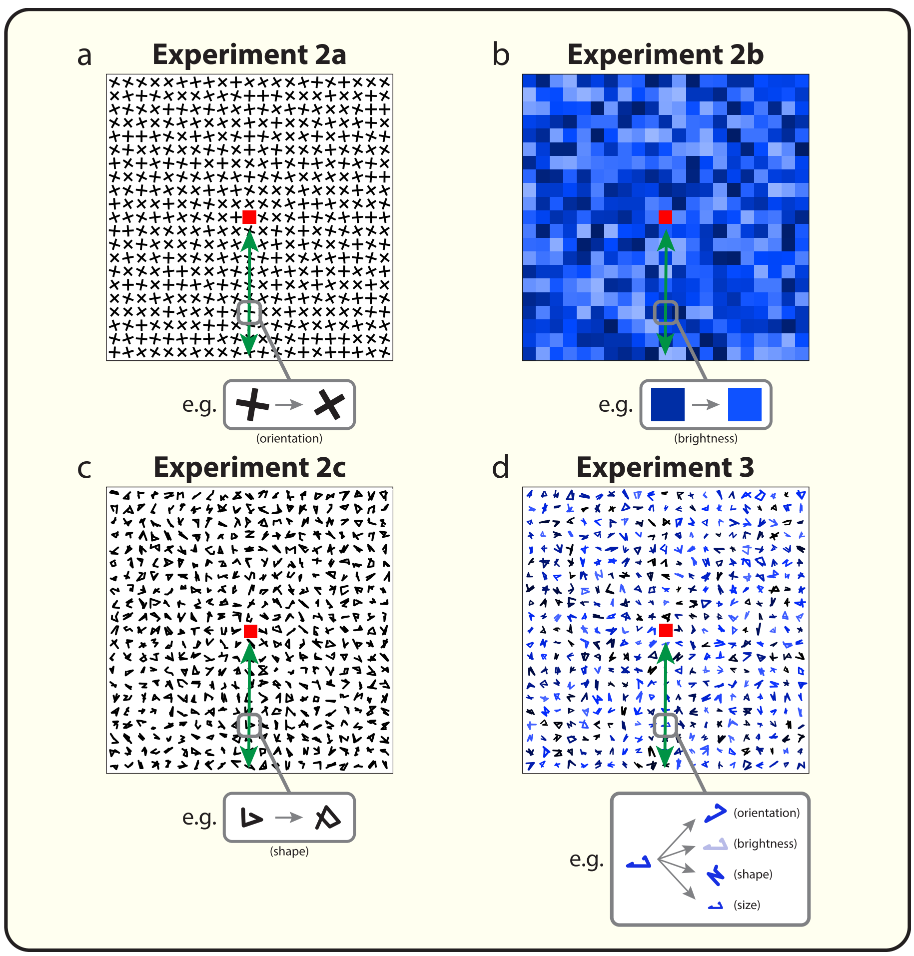

Fig. 3. Depiction of stimuli from Experiments 2–3, with change-defined objects implemented as propagating changes in (a) orientation (Experiment 2a), (b) brightness (Experiment 2b), (c) shape (Experiment 2c), or (d) multiple change types (Experiment 3). In each display, a change-defined object oscillated in a linear path between the central red square and a randomly-chosen edge of the grid (indicated here by the green arrows, which did not appear in the displays themselves), while the paths in the three other cardinal directions were populated by distractors (with elements changing in a random order; not shown). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

noise-dot background (Chubb & Sperling, 1988). But this work showed only that observers are able to identify the dominant fieldwide apparent direction of movement (left vs. right), without any need to track discrete units (i.e. objects). As a result, second-order motion mechanisms may contribute to our displays, but previous results in no way entail the current results, since they need not give rise to (persistent, dynamic, trackable) object representations (Scholl, 2007).

1.3. The current study

Can people perceive and track change-defined objects? We show in four experiments that they can: observers spontaneously describe such stimuli as moving objects (Experiment 1); they can reliably locate a change-defined object among distractors, for various types of change (Experiments 2), even when the type of change itself changes from moment to moment (Experiment 3); and they can even simultaneously track multiple change-defined objects (Experiment 4). Just as important, we also supply animations in which readers can experience each of these effects for themselves, as compelling phenomenological demonstrations (Supplemental Animations S1-S4; see also https://perception.yale.edu/change-tracking/).

2. Experiment 1: Spontaneous phenomenology

We first tested whether change-defined objects are perceived and tracked in perhaps the most direct possible way: we simply showed observers such a display (with orientation changes to crosses; Fig. 2; Animation S1.1) and asked them (without mentioning objects or even motion) to freely describe what they saw. A control group saw a 'spatially scrambled' version of the same display, in which changes occurred at random locations, instead of propagating in a spatiotemporally continuous manner (Animation S1.2).

2.1. Method

The hypotheses, methods, and analyses were preregistered for all experiments prior to data collection (Experiment 1: https://aspredicted.org/CJZ_986 [7.9.24; Data collection began 7.10.24]; Experiment 2a: https://aspredicted.org/B3W_DTG [7.9.24; Data collection began 7.10.24]; Experiment 2b: https://aspredicted.org/S2Y_4B8 [7.9.24; Data collection began 7.10.24]; Experiment 2c: https://aspredicted.org/ZNQ_M4X [7.9.24; Data collection began 7.10.24]; Experiment 3: https://aspredicted.org/LDT_CD6 [7.9.24; Data collection began 7.11.24]; Experiment 4: https://aspredicted.org/BL1_42K [7.9.24; Data collection began 7.11.24]). There were no deviations from the preregistrations for any experiment. Primary data and analysis scripts are available as online Supplemental Information.

2.1.1. Observers

Forty observers participated using the Prolific online platform (Palan & Schitter, 2018) for monetary compensation, with this preregistered sample size determined before data collection began. Observers were excluded (with replacement) according to four preregistered criteria. First, in a post-experimental debriefing phase, observers self-reported how well they paid attention (on a continuous scale, between 1 = "not focused at all", and 100 = "fully focused throughout"), and we excluded observers who self-reported an attention level below 80 ($n = 11$). Second, we excluded any observers (not already excluded by the first criterion) who failed to correctly answer any of three trivial attention check questions (e.g. "How many legs do elephants have? One, nine, or four?"; $n = 1$). Third, we excluded any observers who reported (in response to a debriefing question) that they failed to fixate a central dot, as they were instructed to do ($n = 0$). Fourth, we excluded observers (who were not already excluded via the previous criteria) whose response time was more than 3 median absolute deviations from the median response time of all observers ($n = 3$). All observers gave consent prior to participation, and this research complied with all ethical regulations of (and was approved by) the Yale University Institutional Review Board, and adhered to the Declaration of Helsinki principles and guidelines.

2.1.2. Apparatus

After agreeing to participate, observers were redirected to a website where stimulus presentation and data collection were controlled via custom software written using a combination of HTML, CSS, JavaScript, PHP, and the JsPsych libraries (de Leeuw, 2015). Observers completed the experiment in fullscreen mode on either a laptop or desktop computer.

2.1.3. Stimuli

Observers initially viewed a regular 21 × 21 grid (collectively spanning 60% of the height of the observer's display) of randomly-oriented black crosses (each subtending 85% of its grid cell) on a white background, with the central cross replaced by a red disc (HSB 0/100/100, subtending 25% of its grid cell) for fixation. Changes were implemented (as depicted in Fig. 2) by having one of the crosses change instantly to a new random orientation (differing by 15–75°, while remaining centered in the same grid cell). An initial change occurred to a randomly chosen cross, with subsequent changes occurring every 100 ms. Each subsequent change occurred to one of the 8 crosses immediately adjacent (in cardinal or diagonal directions) to the previous change, with randomly chosen motion directions being updated (by 45° in either direction, chosen randomly) once every 500 to 1500 ms (and with reflection at the display boundaries), for a total animation duration of 10 s. A spatially scrambled control animation was identical except that subsequent changes occurred not to neighboring crosses, but to randomly chosen crosses (constrained to be at least 5 grid cells away from the previous change).

2.1.4. Procedure

Observers were instructed to pay attention carefully to a single animation while fixating the central red dot. After the 10s animation concluded, they were prompted: "Describe what you saw. Please answer in a single sentence (using roughly the amount of space in the box)", where the box had space for about 100 characters. 20 observers each viewed either the intact or spatially scrambled animation.

2.1.5. Analyses

Observers' responses were categorized by two independent coders (one the first author, the other naïve), who were blind to the condition to which the observers were assigned. The coders categorized the responses into two categories, based on whether or not the responses described a single moving object (with the exact coder instructions provided in the preregistration). The inter-coder reliability was 97.5% — with the coders' categorization agreeing on all forty responses except one (and this single response was categorized as not reflecting a moving object).

2.2. Results and discussion

A majority of observers (65%) spontaneously described the change-defined object display in terms of a single moving object, writing for example that "I saw a cross moving around the screen", or "an object was moving through spikes making them move". In contrast, not a single observer did so for the spatially scrambled control display, writing instead for example of "+'s changing into x's and x's changing into +'s", or "flashing crosses across the screen" — a highly reliable difference ($\chi^2(1, N = 40) = 19.26, p < 0.001, d = 1.93$, 95% CI [1.07, 2.79]). This thus constitutes an initial demonstration of how change-defined objects are spontaneously perceived and tracked.

3. Experiment 2a: Localizing objects

We next demonstrated the perception of change-defined objects using an objective localization measure. Observers viewed similar displays in which a change-defined object (implemented via orientation changes) oscillated back and forth between the center of the display and one of its four sides (Fig. 3a; Animation S2.1), while spatially-scrambled changes were presented in the other three directions. Observers simply had to locate the change-defined object.

3.1. Method

This experiment was identical to Experiment 1 except as noted.

3.1.1. Observers

Twenty new observers participated, with this preregistered sample size determined before data collection began. Observers were excluded (with replacement, and again according to the preregistration) according to the first two criteria from Experiment 1 (here excluding 3 [self-reported focus] and 0 [attention-check questions], respectively). In addition, we excluded observers whose accuracy was more than three standard deviations away from the group mean ($n = 0$).

3.1.2. Stimuli

The collective grid was now sized to exactly 600 × 600 px, with individual crosses relatively sized as in Experiment 1. The central element was a red square subtending its entire grid cell. To the right of the grid were virtual buttons with arrows (up, down, left, right), above which was written: "Where is the moving object relative to the red square? Click on the appropriate button below:". The first cross to change was always adjacent to the fixation square, in one of the four cardinal directions. And subsequent changes occurred along that same cardinal path (as depicted in the upper left quadrant of Fig. 3), with the trajectory eventually reflecting off the grid boundary and returning to the fixation square, after which the sequence repeated — thus yielding a change-defined object that oscillated back and forth between the fixation square and the display edge. Changes also occurred simultaneously in the other three cardinal directions, but these were spatially scrambled, such that the next change always occurred to a randomly chosen cross along that cardinal direction (with subsequent changes always at least 5 cells apart).

3.1.3. Procedure

Observers were instructed to click on the relevant virtual arrow button to indicate on each trial which cardinal direction contained the moving object, with each observer completing 8 trials (2 with a target in each cardinal direction, presented in a random order).

3.2. Results and discussion

Observers were highly accurate, and well above 25%-chance (M = 81.3%, $SD = 0.35$, $t(19) = 7.28$, $p < 0.001$, $d = 1.63$, 95% CI [0.55, 2.71]). And as an additional indication of how relatively easy it was to see the change-defined objects, we note that 65% of the observers performed at 100% accuracy across all trials. Beyond the subjective reports in Experiment 1, this shows how the perception and tracking of change-defined objects also manifests in terms of objective performance.

4. Experiments 2b and 2c: Beyond orientation

Is the perception of change-defined objects in some way specific to orientation changes? To find out, we sought to replicate the effect from Experiment 2a with changes in brightness (Experiment 2b; Fig. 3b; Animation S2.2) or shape (Experiment 2c; Fig. 3c; Animation S2.3).

4.1. Method

These experiments were identical to Experiment 2a except as noted.

4.1.1. Observers

Twenty new observers participated in each experiment, with this preregistered sample size determined before data collection began to exactly match that of Experiment 2a. Observers were excluded (with replacement) via the same criteria for each of Experiments 2b and 2c (self-reported focus: $ns = 5$ and 2, respectively; attention-check questions: $ns = 0$ and 0; accuracy outliers: $ns = 1$ and 0).

4.1.2. Stimuli

In Experiment 2b, each element was a colored square that filled its grid cell, with a hue value of 223, a saturation value of 100%, and a random brightness value between 20 and 80 (as depicted in Fig. 3b). Each change was implemented by changing the relevant cell's brightness value by a random amount (of at least 20, while remaining in the range of 20–80). In Experiment 2c, each element was a black shape generated by four random connected strokes (each at least 25.5% of the cell width, without extending past the cell boundary; as depicted in Fig. 3c). Each change was implemented by changing the relevant cell's shape to a different random shape.

4.2. Results and discussion

Observers were again highly accurate, and well above chance (25%), both for brightness changes ($M = 95.6\%$, $SD = 0.12$, $t(19) = 27.07$, $p <$ 0.001, $d = 6.05$, 95% CI [3.84, 8.27]), and shape changes ($M = 84.4\%$, $SD = 0.29$, $t(19) = 9.18$, $p < 0.001$, $d = 2.05$, 95% CI [0.90, 3.21]). And again, a majority of observers (85% and 75%, respectively) performed at 100% accuracy. These results show that the perception and tracking of change-defined objects is a general phenomenon, having no specific connection to orientation changes (as explored in Experiments 1 and 2a), per se.

5. Experiment 3: Variable types of changes

In the previous displays, each change always involved the same dimension (orientation, brightness, or shape). We next tested whether this regularity was necessary: observers localized change-defined objects involving different *types* of changes from moment to moment — as in an orientation change followed in the next frame by a shape change, followed by a brightness change, etc., with no neighboring changes ever involving the same type (Fig. 3d; Animations S3.1 and S3.2). (This sort of manipulation has never to our knowledge been previously studied, even in the literature on second-order motion.)

5.1. Method

This experiment was identical to Experiment 2a except as noted.

5.1.1. Observers

Twenty new observers participated, with this preregistered sample size determined before data collection began to exactly match that of Experiments 2a-2c. Observers were excluded (with replacement) via the same criteria (self-reported focus: $n = 4$; attention-check questions: $n =$ 0; accuracy outliers: $n = 0$).

5.1.2. Stimuli

As depicted in Fig. 3d, each element was a random shape (generated as in Experiment 2c), with a random brightness level (chosen as in Experiment 2b, in the range of 0–65), and presented with a random size (between 60% and 100% of the full size from Experiment 2c). Each individual change then involved a randomly chosen change type (orientation, brightness, shape, or size), with the other three properties conserved. Moreover, neighboring changes always involved different change types. (Changes to orientation, brightness, and shape were implemented as in Experiments 2a-2c. Size changes involved choosing a new random size between 60% and 100% of the full size, with a minimum change of 15%).

5.2. Results and discussion

Performance was again far above chance ($M = 72.5\%$, $SD = 0.35$, $t(19) = 6.06$, $p < 0.001$, $d = 1.36$, 95% CI [0.32, 2.39]), and 55% of observers performed perfectly across all trials. In many ways, this is the most extreme example of change-defined objects that we have seen so far, in which perception and tracking are possible even when the change type itself is in constant flux.

6. Experiment 4: Multiple object tracking

Observers can spontaneously perceive and localize individual change-defined objects, but can they also track *multiple* objects at once — keeping them distinct from other identical objects over time and motion? We explored this using what may be the hallmark task of such abilities: multiple-object tracking (MOT; Holcombe, 2023; Pylyshyn & Storm, 1988; Scholl, 2009). In MOT, observers first see several objects, a subset of which are flashed as targets. All of the (again-identical) objects then move on unpredictable trajectories. Afterwards, observers simply indicate which objects (now in new positions) were the initially-flashed targets. Is the representation of change-defined objects robust enough to support MOT? We investigated this by asking observers to track two out of four simultaneously moving change-defined objects (via orientation changes; Fig. 4; Animation S4).

6.1. Method

This experiment was identical to Experiment 1 except as noted.

6.1.1. Observers

Twenty new observers participated, with this preregistered sample size determined before data collection began to exactly match that of Experiments 2a-2c and 3. Per the preregistration, observers were excluded (with replacement) via the same first three criteria from Experiment 1, along with the accuracy-outlier criterion from Experiments 2a-2c (self-reported focus: $n = 4$; attention-check questions: $n = 0$; fixation compliance: $n = 0$; accuracy outliers: $n = 0$).

6.1.2. Stimuli

Observers viewed a 31×31 grid (collectively sized to 800×800 px), with the central element being a green (HSB 120/100/100) fixation dot (sized to 85% of its grid cell). Four different change-defined objects simultaneously moved about the display in the same manner as Experiment 1, with different initial starting positions (randomly determined, but with one within each of the four grid quadrants) and motion directions. Whenever two objects became closer than 6 grid cells, they repelled each other (implemented by immediately choosing opposing motion directions for each).

6.1.3. Procedure

At the beginning of each trial, the crosses corresponding to two of the change-defined objects (chosen randomly) were flashed for 2500 ms in red to indicate their status as targets, while the other two flashed in blue (to indicate their status as distractors), and the subsequent motion (with all crosses again drawn in black) lasted 7.5 s, after which the crosses corresponding to the four objects each turned into four red question marks. Observers were instructed to fixate the central disc during the motion, and to click in the final tableau on the two question-marks corresponding to the initial red targets. This task could only be completed by tracking the two targets as the same persisting individuals throughout the motion period on the basis of their history.

6.2. Results and discussion

Multiple-object tracking with change-defined objects was successful ($M = 88.8\%$, $SD = 0.17$) — far above 50%-chance ($t(19) = 10.33$, $p < 0.001$, $d = 2.31$, 95% CI [1.10, 3.52]), and similar to typical MOT performance (Holcombe, 2023; Pylyshyn & Storm, 1988; Scholl, 2009). To our knowledge, this type of explicit object tracking ability has never before been demonstrated with any phenomenon of second-order motion.

7. General discussion

The central message of this study is that people can readily perceive and track change-defined objects — as measured by spontaneous phenomenological reports (Experiment 1), localization (Experiments 2 and 3), and even multiple-object tracking (Experiment 4).

7.1. It's (always) time for a (pure) change

On their face, these results may seem almost paradoxical: people reliably perceived and tracked moving objects even when there was *nothing to track*, since no features (or even change types, in Experiment 3) persisted from frame to frame; and in static frames it is impossible even in principle to segment change-defined objects from the background. As such, these results show that dynamic object perception need not involve featural identification and segmentation; instead, dynamic objects can be perceived based on the spatiotemporal continuity of *pure change* transients themselves.

The brute detection of motion under such circumstances should not be surprising, since these manipulations are based in part on well-characterized mechanisms of second-order motion and apparent motion (Nishida, 2011). Instead, the key contribution of these phenomena is to draw a connection for the first time between (1) second-order motion perception, and (2) the ability to perceive and track individual objects as the same persisting entities over time (Scholl, 2007). Over multiple decades, these two prominent domains of investigation have progressed almost entirely independently, with neither typically engaging with (or even citing) the other. As such, the current phenomena aim to join these two literatures.

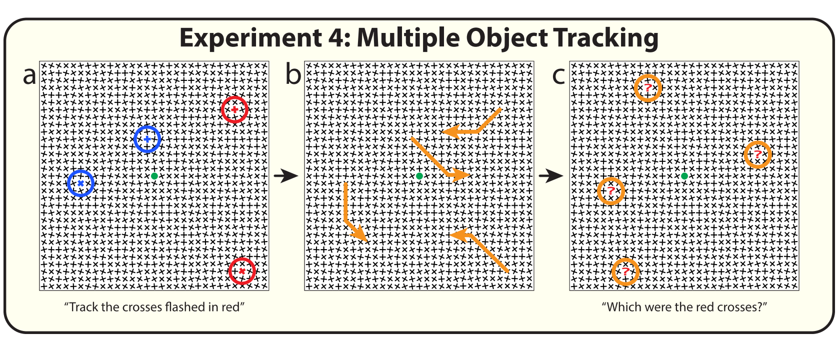

Fig. 4. Depiction of the multiple-object tracking (MOT) task from Experiment 4, in which observers had to track two out of four change-defined objects. At the beginning of each trial, two target crosses flashed in red (a), while two other distractor crosses flashed in blue. All four then turned black (identical to the background texture), and started moving around the grid (b). After several seconds, all four objects stopped and turned into question marks, and observers clicked on the two that they thought were the targets (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7.2. Objects, features, grouping, events, and changes

Change-defined objects also seem theoretical rich insofar as they directly engage with many other prominent themes from vision research and cognitive science (even beyond those already discussed in Introduction). Below we briefly discuss several examples.

The contribution of features to object tracking has long been a central topic in cognitive science (Yu & Lau, 2023), especially in the context of the 'binding problem' (Roskies, 1999; Treisman, 1996, 1998): given that some features (e.g. color and shape) are processed via distinct circuits, how and when do those free-floating properties get 'reassembled' into cohesive object representations (e.g. so that a red square and a blue triangle can be distinguished from a red triangle and a blue square)? An influential class of 'pointer'-based accounts posits that objects are bound to spatial locations, in a way that is independent of any particular surface features (Awh & Vogel, 2025; Kahneman et al., 1992; Pylyshyn, 1989). The present results are in line with such views, since transient changes may still function as reliable indicators of changing object locations over time. Surface features, then, may be assigned to objects via a secondary process (which may also sometimes go awry, as in illusory conjunctions; Treisman & Schmidt, 1982), but they do not define the objects in the first place. In our study, this dynamic may be especially clear from Experiment 3, in which observers reliably perceived persisting objects even though any specific feature value (of orientation, shape, color, or size) appeared only momentarily. This could also be related to the perception of grouping based on common fate (Wagemans et al., 2012; Wertheimer, 1923) — insofar as that 'fate' is determined by spatiotemporal factors (Alais et al., 1998; Lee & Blake, 1999) rather than surface features. But of course the individual elements in such displays typically maintain the same surface features over time, unlike the change-defined objects explored here.

Change-defined objects may also relate to *events*. Indeed, each individual transient may itself intuitively constitute an event in some sense. We nevertheless think that our displays are best described as change-defined *objects*, for three related reasons. First, doing otherwise risks failing to distinguish an object itself from its properties. Although redness may be part of the percept a moving red disc, that object does not *reduce* to redness, and one might continue to track "the same object" even it suddenly turns blue; similarly, although a transient event may be a momentary part of the percept of a moving change-defined object, that object does not *reduce* to the transients, and one might continue to track "the same object" even if the underlying types of transients suddenly change (e.g. from luminance changes to orientation changes). (This is in part why observers in Experiments 2 and 3 saw only a single object per display — even though there were equated transients occurring in all four cardinal directions.) Second, the vast majority of observers in Experiment 1 clearly limned their experience in terms of singular objects (e.g. referring to "A moving X", or "I could see a little black cross moving around the area"), rather than events. Third, we worry that describing change-defined objects in terms of "events" could be sociologically awkward — since other work on second-order motion perception has not previously been described in that way.

Few studies have explored how featural changes themselves may *contribute* to object tracking (even though many have investigated how changes may *disrupt* object tracking, for example in the tunnel effect). The only two exceptions we know of focused on how changes may help to distinguish two otherwise similar-looking objects. In one study, when viewing two superimposed stationary Gabor patches that underwent continuous featural changes (progressively shifting in color and orientation), observers successfully tracked one over time (Blaser et al., 2000). Similarly, another study showed that 'object-specific preview benefits' — a hallmark of automatic object tracking (Kahneman et al., 1992) — applied not only to moving objects with persisting properties but also to objects constantly changing features (e.g. a disc continuously shifting color; Mazalik et al., 2023). These two studies show respectively that people can track through feature space in stationary objects, and that change itself can serve as a distinctive feature for tracking. But the changes in these studies were *continuous* and thus *predictable* (and this is of course how many real-world objects behave, as when the shape of a tiger's body changes gradually as it walks). In contrast, the perturbations in change-defined objects are maximally unpredictable — such that their future appearance can never be predicted from the current frame.

7.3. Conclusions: Seeing into the tall grass?

Why might the visual system so readily construct persisting object representations out of only spatiotemporal properties? After all, most objects in our visual world do not change their appearance randomly from moment to moment! This mechanism might nevertheless allow us to perceive dynamic objects in (rare but important) real-world situations not on the basis of what they look like, but only on the basis of their behavior, even when they are themselves invisible — as when a moving tiger hidden in the tall grass can nevertheless be tracked via the (ever-changing) perturbations of the surrounding grass itself.

Studies in Human

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place.

The Yale University Institutional Review Board granted an exemption. The authors provided the following explanation: N/A.

CRediT authorship contribution statement

Dawei Bai: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Brian J. Scholl:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Conceptualization.

Informed consent and patient details

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

Funding

This research received no specific funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

For helpful discussion and/or feedback on previous drafts, we thank Ed Awh, Mario Belledonne, Henry Jones, Peter Mazalik, Sam McDougle, Zekun Sun, Johan Wagemans, and the members of the Yale Perception & Cognition Laboratory.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2026.106657.

Data availability

The raw data for all studies are included in the Supplementary Material file.

References

Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience, 1*(2), 160–164. https://doi.org/10.1038/414

Awh, E., & Vogel, E. K. (2025). Working memory needs pointers. *Trends in Cognitive Sciences, 29*(3), 230–241. https://doi.org/10.1016/j.tics.2024.12.006

Badcock, D. R., & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research, 25*(9), 1253–1258. https://doi.org/10.1016/0042-6989(85)90040-9

Bahill, A. T., & LaRitz, T. (1984). Why can't batters keep their eyes on the ball? *American Scientist, 72*(3), 249–253.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature, 408*(6809), 196–199. https://doi.org/10.1038/35041567

Bower, T. G. R., Broughton, J., & Moore, M. K. (1971). Development of the object concept as manifested in changes in the tracking behavior of infants between 7 and 20 weeks of age. *Journal of Experimental Child Psychology, 11*(2), 182–193. https://doi.org/10.1016/0022-0965(71)90074-9

Bremner, J. G., Johnson, S. P., Slater, A., Mason, U., Foster, K., Cheshire, A., & Spring, J. (2005). Conditions for young infants' perception of object trajectories. *Child Development, 76*(5), 1029–1043. https://doi.org/10.1111/j.1467-8624.2005.00895.x

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology, 4*, 121–138. https://doi.org/10.1080/17470215208416611

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4*(2–3), 103–129. https://doi.org/10.1163/156856889X00077

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*(11), 1986–2007. https://doi.org/10.1364/JOSAA.5.001986

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Flombaum, J. I., Kundey, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in rhesus macaques: A study of the tunnel effect. *Psychological Science, 15*(12), 795–800. https://doi.org/10.1111/j.0956-7976.2004.00758.x

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 840–853. https://doi.org/10.1037/0096-1523.32.4.840

Flombaum, J. I., Scholl, B. J., & Pylyshyn, Z. W. (2008). Attentional resources in visual tracking through occlusion: The high-beams effect. *Cognition, 107*(3), 904–931. https://doi.org/10.1016/j.cognition.2007.12.015

Gao, T., & Scholl, B. J. (2010). Are objects required for object files?: Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition, 18*(1), 82–109. https://doi.org/10.1080/13506280802614966

Holcombe, A. (2023). *Attending to moving objects*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009003414

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences, 100*(18), 10568–10573. https://doi.org/10.1073/pnas.1630655100

Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions on Information Theory, 8*(2), 84–92.

Julesz, B., & Bergen, J. R. (1983). Human factors and behavioral science: Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal, 62*(6), 1619–1645. https://doi.org/10.1002/j.1538-7305.1983.tb03502.x

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Koerfer, K., & Lappe, M. (2022). Perceived movement of nonrigid motion patterns. *PNAS Nexus, 1*(3), Article pgac088. https://doi.org/10.1093/pnasnexus/pgac088

Koerfer, K., Watson, T., & Lappe, M. (2024). Inability to pursue nonrigid motion produces instability of spatial perception. *Science Advances, 10*(45), Article eadp6204. https://doi.org/10.1126/sciadv.adp6204

Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science, 284*(5417), 1165–1168. https://doi.org/10.1126/science.284.5417.1165

Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T.-K. (2021). Multiple object tracking: A literature review. *Artificial Intelligence, 293*, Article 103448. https://doi.org/10.1016/j.artint.2020.103448

Mather, G., & Anstis, S. (1995). Second-order texture contrast resolves ambiguous apparent motion. *Perception, 24*(12), 1373–1382. https://doi.org/10.1068/p241373

Mazalik, P., Wu, Q., Flombaum, J., & Halberda, J. (2023). The more things change the more they stay the same; a continuously changing item can define a visual object [abstract]. *Journal of Vision, 23*(9), 5186. https://doi.org/10.1167/jov.23.9.5186

McLeod, P. (1987). Visual reaction time and high-speed ball games. *Perception, 16*(1), 49–59. https://doi.org/10.1068/p160049

Milan, A., Leal-Taixe, L., Reid, I., Roth, S., & Schindler, K. (2016). MOT16: A benchmark for multi-object tracking. *arXiv*, 10.48550/arXiv.1603.00831.

Nishida, S. Y. (2011). Advancement of motion psychophysics: Review 2001–2010. *Journal of Vision, 11*(5), 1–53. https://doi.org/10.1167/11.5.11

Nygård, G. E., Sassi, M., & Wagemans, J. (2011). The influence of orientation and contrast flicker on contour saliency of outlines of everyday objects. *Vision Research, 51*(1), 65–73. https://doi.org/10.1016/j.visres.2010.09.032

Nygård, G. E., Van Looy, T., & Wagemans, J. (2009). The influence of orientation jitter and motion on contour saliency and object identification. *Vision Research, 49*(20), 2475–2484. https://doi.org/10.1016/j.visres.2009.08.002

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65–97. https://doi.org/10.1016/0010-0277(89)90014-0

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197. https://doi.org/10.1163/156856888X00122

Regan, D., & Beverley, K. I. (1984). Figure-ground segregation by motion contrast and by luminance contrast. *Journal of the Optical Society of America A, 1*(5), 433–442.

Regan, D., & Hamstra, S. (1991). Shape discrimination for motion-defined and contrast-defined form: Squareness is special. *Perception, 20*(3), 315–336. https://doi.org/10.1068/p200315

Reichardt, W., & Poggio, T. (1979). Figure-ground discrimination by relative movement in the visual system of the fly: Part I: Experimental results. *Biological Cybernetics, 35*(2), 81–100. https://doi.org/10.1007/BF00337434

Roskies, A. L. (1999). The binding problem. *Neuron, 24*(1), 7–9. https://doi.org/10.1016/S0896-6273(00)80817-X

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language, 22*(5), 563–591. https://doi.org/10.1111/j.1468-0017.2007.00321.x

Scholl, B. J. (2009). What have we learned about attention from multiple-object tracking (and vice versa)? In D. Dedrick, & L. Trick (Eds.), *Computation, cognition, and Pylyshyn* (pp. 49–78). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/8135.003.0005.

Tavassoli, A., & Ringach, D. L. (2010). When your eyes see more than you do. *Current Biology, 20*(3), R93–R94. https://doi.org/10.1016/j.cub.2009.11.048

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178. https://doi.org/10.1016/S0959-4388(96)80070-5

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1295–1306. https://doi.org/10.1098/rstb.1998.0284

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14*(1), 107–141. https://doi.org/10.1016/0010-0285(82)90006-8

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin, 138*(6), 1218–1252. https://doi.org/10.1037/a0029334

Wertheimer, M. (1923). Untersuchungen zur Lehre von der gestalt, II [investigations in gestalt theory, II: Laws of organization in perceptual forms]. *Psychologische Forschung, 4*, 301–350.

Yu, X., & Lau, E. (2023). The binding problem 2.0: Beyond perceptual features. *Cognitive Science, 47*(2), Article e13244. https://doi.org/10.1111/cogs.13244

Zhang, L., Li, Y., & Nevatia, R. (2008). Global data association for multi-object tracking using network flows. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). https://doi.org/10.1109/CVPR.2008.4587584