

# The what–where trade-off in multiple-identity tracking

Michael A. Cohen · Yair Pinto · Piers D. L. Howe ·  
Todd S. Horowitz

Published online: 5 March 2011  
© Psychonomic Society, Inc. 2011

**Abstract** Observers are poor at reporting the identities of objects that they have successfully tracked (Pylyshyn, *Visual Cognition*, 11, 801–822, 2004; Scholl & Pylyshyn, *Cognitive Psychology*, 38, 259–290, 1999). Consequently, it has been claimed that objects are tracked in a manner that does not encode their identities (Pylyshyn, 2004). Here, we present evidence that disputes this claim. In a series of experiments, we show that attempting to track the identities of objects can decrease an observer’s ability to track the objects’ locations. This indicates that the mechanisms that track, respectively, the locations and identities of objects draw upon a common resource. Furthermore, we show that this common resource can be voluntarily distributed between the two mechanisms. This is clear evidence that the location- and identity-tracking mechanisms are not entirely dissociable.

**Keywords** Multiple object tracking · Attention · Memory

When we interact with a dynamic environment, such as in a soccer game or a crowded playground, we often need to keep track of the positions of discrete individuals. An extensive literature on multiple-object tracking (MOT; Pylyshyn &

Storm, 1988) has demonstrated that humans can track the positions of multiple independently moving objects (for reviews, see Cavanagh & Alvarez, 2005; Scholl, 2009). The moving objects in MOT studies have typically been identical, so as to prevent observers from using featural cues to aid tracking. However, in everyday life, tracked objects usually have distinct identities. Recent work by Makovski and colleagues (Makovski & Jiang, 2009; Makovski, Vazquez, & Jiang, 2008) has demonstrated that objects with unique features are easier to track (see also Horowitz et al., 2007). However, there is an important difference between being able to use feature information to improve tracking and being able to track the identities of distinct objects. For example, Yantis (1992) asked observers to track targets that moved at a common speed, which differed from the common speed of the distractors. Even though observers were at chance at discriminating between the two speeds, they could still use the speed difference to help segregate targets from distractors. While some studies have indicated that observers have very little conscious access to featural or identity information about the tracked targets (Pylyshyn, 2004), it has been shown that observers can report both features (Howard & Holcombe, 2008) and identities of at least some of the tracked objects (Horowitz et al., 2007; Oksama & Hyönä, 2004, 2008). Here we asked whether position tracking and identity tracking share a common resource.

A few studies of tracking have employed unique targets, asking observers to report which target is at which position (Horowitz et al., 2007; Oksama & Hyönä, 2004; see also Oksama & Hyönä, 2008; Pylyshyn, 2004). Following Oksama and Hyönä (2004), we call this task “multiple-identity tracking,” or MIT, to distinguish it from standard MOT, in which the features or identity of the stimuli are irrelevant to the task (that is, typical MOT with identical objects, as well as the experiments of Makovski & Jiang, 2009). In Pylyshyn (2004), targets were individuated by either starting position (targets originated in different corners of the display) or by a digit label (presented during

---

M. A. Cohen (✉)  
Department of Psychology, Harvard University,  
33 Kirkland St., 760 William James Hall,  
Cambridge, MA 02138, Canada  
e-mail: michaelthecohen@gmail.com

Y. Pinto  
University of Amsterdam,  
Amsterdam, Netherlands

P. D. L. Howe  
University of Melbourne,  
Melbourne, Australia

T. S. Horowitz  
Brigham and Women’s Hospital, Harvard Medical School,  
Boston, MA, USA

the target designation phase). Critically, the feature that identified the targets (e.g., the digit) was not visible during the motion phase. In contrast, Oksama and Hyönä (2004) and Horowitz et al. (2007) used distinct objects (or pseudo-objects) that were continuously visible throughout the tracking period and masked only in the testing period. Although unique objects are easier to track than identical objects, both studies found that the capacity for identities was systematically lower than the capacity for locations. That is, observers could successfully report the location of a target while being unable to report its identity.

These studies suggest that there is a difficulty to binding both locations and identities in MOT (Imaruoka, Saiki, & Miyauchi, 2005; Saiki, 2002). Interestingly, observers were better at maintaining the identities of tracked items in the Horowitz et al. (2007) and Oksama and Hyönä (2004) studies than in the Pylyshyn (2004) study. This suggests that the problem of binding locations to identities can be partially solved by keeping the identities of the targets visible throughout the tracking period.

Even with continuously visible stimuli, not every tracked item is bound to its identity (Horowitz et al., 2007; Oksama & Hyönä, 2004, 2008). Horowitz et al. suggested that this phenomenon could be explained if there were two independent systems responsible for tracking positions and identities, respectively. For example, Pylyshyn (1989, 2001, 2007) has proposed a FINST (Fingers of INSTantiation) model of spatial cognition, in which MOT is mediated by early, preattentive, pointer-like indexes that pick out a small number of objects to be tracked. These pointers are thought to only encode the spatiotemporal properties of objects and not to represent features like shape or color. Binding features to locations, on the other hand, seems to require focused attention (Wheeler & Treisman, 2002). Thus, a preattentive system of indexes would track positions, whereas selective attention would track identities, but with a lower capacity, because it is more resource intensive. We shall call this the *independent mechanisms model*.

The other obvious possibility, which we will refer to as the *common resource model*, is that the position- and identity-tracking mechanisms share a common resource that can be flexibly distributed between the two mechanisms. For example, Zhang and Luck (2008) proposed a model of visual working memory in which there are a fixed number of slots and a limited resource that can be flexibly divided up so that each slot has a different resolution. Since working memory seems tightly connected to MOT (Allen, McGeorge, Pearson, & Milner, 2006; Fougny & Marois, 2006; Howe, Horowitz, Morocz, Wolfe, & Livingstone, 2009; Oksama & Hyönä, 2004), it seems natural to extend such a model to the MIT domain. In this case, we could explain the lower capacity for identities by suggesting that encoding item positions with sufficient resolution to distin-

guish targets from distractors requires fewer resources than encoding item identities with sufficient resolution to successfully distinguish which target was which on all trials.

How can we distinguish these two models? The independent mechanisms model assumes that the relative capacity limits are a fixed property of each system, so that, for example, making the identity task harder should not affect the location task. In contrast, the common resource model assumes that any additional resources required by the identity task would come at the expense of reduced performance on the location task. The simplest test would be to simply compare location performance during MOT and MIT. The common resource model predicts that asking people to encode identities while tracking should reduce overall tracking performance. Both Pylyshyn (2004) and Horowitz et al. (2007) tested this prediction by comparing dual-task to single-task performance, and both found that observers were equally good at reporting target locations when asked to track identities and when asked to only track locations. Horowitz et al. took this as evidence in favor of the independent mechanisms model.

However, there are problems with both experiments. As noted above, identities were not visible during the tracking phase of Pylyshyn's (2004) experiments. It could be that the identities are only available during the cuing phase, the task becomes so difficult that observers largely abandon trying to track target identities and instead focus on the location task. This would explain why observers performed so poorly on the identity task. Horowitz et al. (2007) solved this problem by allowing object identities to be continuously visible. However, they compared dual- and single-task performance across different groups of observers. Moreover, they did not manipulate task difficulty aside from in the dual-task manipulation itself, in contrast to the Pylyshyn (2004) experiments, which varied tracking speed. It is possible that Horowitz et al.'s observers were operating in a domain where there were sufficient resources to accommodate both tasks. The goal of the experiments reported here was to push observers into a domain where we might be able to observe the trade-off predicted by the common resource model. Thus, we combined Pylyshyn's (2004) design with Horowitz et al.'s stimuli.

To preview our results, in [Experiments 1–3](#) we found that location performance dropped significantly when observers were asked to maintain target identities, as compared to pure location tracking. This showed that the mechanisms that track, respectively, the locations and the identities of objects share a common resource. [Experiments 4–6](#) showed that this common resource could be voluntarily distributed between the two mechanisms. Taken together, these results suggest that the independent mechanisms model is incorrect: Location tracking is not carried out independently of identity tracking.

## Experiment 1

Our first experiment was modeled after Pylyshyn (2004), except that object identities remained visible during tracking (Horowitz et al., 2007; Oksama & Hyönä, 2004, 2008). Thus, observers tracked cartoon animals that were visible throughout the tracking epoch. The stimuli were then masked at the response phase. In the *location* condition, observers were asked to click on the locations of all targets. In the *identity* condition, observers were asked to click on each target in turn as it was named. In the identity condition, we can score performance in two ways: We can determine whether the participants correctly matched animals with positions (*identity accuracy*), or we can determine how many selected locations were target locations, regardless of the correct placement of identities (*location accuracy*). Note that under these definitions, identity accuracy can never exceed location accuracy but can be equal to it. As in the Pylyshyn study, we also varied the speed of the objects, on the theory that faster targets will be more difficult to track (Alvarez & Franconeri, 2007), at least in part because they cover a greater distance (Franconeri, Jonathan, & Scimeca, 2010), and this might unmask trade-offs between location and identity information.

### Method

**Participants** Eight participants were recruited from the Brigham and Women's Hospital Visual Attention Laboratory volunteer pool. All had either normal or corrected-to-normal visual acuity and had passed the Ishihara color screen. Participants gave informed consent, as approved by the Brigham and Women's Hospital IRB, and were paid \$10/h for participation. Participants ranged in age from 18 to 50 years ( $M = 31.3$  years,  $SD = 12.2$ ).

**Apparatus and stimuli** Stimuli were presented on 21-in. color CRT monitors (Hitachi RasterOps SuperScan Mc801 and Mitsubishi Diamond Pro 91TXM) controlled by Macintosh G5 computers running Mac OS 10.4. All experiments were programmed in MATLAB 7.5 (The MathWorks, Natick, MA) using the Psychophysics Toolbox routines (Brainard, 1997; Pelli, 1997). Monitor spatial resolution was set to  $1,024 \times 768$  pixels with a refresh rate of 75 Hz (13 ms per frame). Observers were seated 57 cm from the monitor; at this distance, 1 cm on the screen subtended  $1^\circ$  of visual angle. The items were selected at random from a set of 22 cartoon animals (see the Appendix), each of which fit within a  $2.37^\circ \times 2.37^\circ$  bounding box (Fig. 1). The background was white.

**Procedure** The experiment consisted of eight blocks of 40 trials. On each trial, eight cartoon animals were randomly

placed on the screen (Fig. 1). At the start of the trial, four of the animals were highlighted by red boxes for 3 s to indicate that these were the targets to be tracked. All items then moved for a random duration uniformly distributed between 3 and 7 s, in increments of 200 ms (i.e., 3,000, 3,200, 3,400, . . . 6,800, 7,000 ms). The length of the trial was unpredictable in order to ensure that observers were actively tracking rather than simply locating the targets just before the end of the trial. All items moved in straight lines, except when they bounced off each other or the sides of an imaginary window ( $32.7^\circ \times 22.5^\circ$ ). The speed of the items varied from block to block. There were four possible speeds:  $0^\circ/s$  (i.e., stationary),  $6.85^\circ/s$ ,  $13.7^\circ/s$ , and  $20.55^\circ/s$ . At the end of a trial, each animal was masked by a  $2.37^\circ \times 2.37^\circ$  red square.

There were two response conditions, presented in separate blocks. In the *location* condition, observers tracked the target animals but did not have to report their identity. After the animals were occluded, observers were instructed to click on the location of each target. In the *identity* condition, observers had to report both locations and identities. After the animals were occluded, observers were given the name of each target in turn and asked to click on its location (e.g., "Please click on the Elephant," "Please click on the Gorilla," etc.). For this condition, two types of data were extracted: location accuracy and identity accuracy. *Location accuracy* refers to participants' ability to differentiate target from distractor locations, while *identity accuracy* refers to their ability to know which specific target is in which specific location. Imagine that a participant were tracking a gorilla, giraffe, crocodile, and kangaroo. If, when asked to click on the kangaroo, he instead clicked on the gorilla, this would be considered a hit with respect to location accuracy, since it indicated that he knew it was a target item, but a miss with respect to identity accuracy, since he did not know which target it was. Note that identity accuracy can never exceed location accuracy under this method. Accuracy in this case was calculated by measuring the percentage of objects that were correctly identified or located.

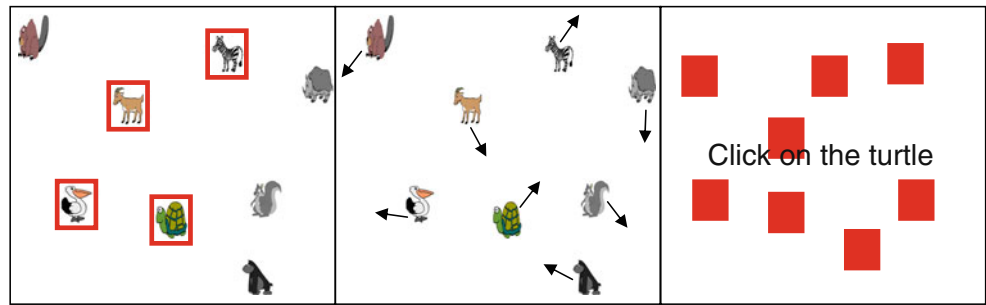
In the figures, we denote the various dependent measures with an abbreviation in the form of *measure/condition*. Location accuracy in the location condition is thus labeled "Loc/Loc," location accuracy in the identity condition is "Loc/ID," and identity accuracy in the identity condition is "ID/ID."

In all cases, feedback was provided after four responses were given. The order in which observers completed the experiment was counterbalanced in terms of the location and identity conditions and the speeds at which the items moved.

### Results and discussion

Accuracy is plotted as a function of target speed in Fig. 2. Performing a two-way within-subjects ANOVA

**Fig. 1** An example of the task we used in [Experiment 1](#). The targets were indicated with red boxes at the beginning of the trial and remained visible throughout the movement phase. When the movement phase stopped, observers would click on the targets, each of which was now covered with a red square



with Speed and Conditions (Loc/Loc, Loc/ID, ID/ID) as factors, we found that overall location accuracy decreased as a function of speed [ $F(3, 21) = 55.7, p < .001$ ]. Our particular interest was in how tracking identities affected location tracking. We found that location tracking was significantly worse when also having to track identities, as compared to when tracking locations only [ $F(1, 7) = 48.8, p < .001$ ]. These two conditions and the speed of the objects interacted, such that the effect was larger as objects increased in speed [ $F(3, 21) = 7.1, p < .01$ ]. Post-hoc tests (Fisher's least significant difference) confirmed that there was no difference between the two conditions in the stationary and slow conditions, but there were significant differences at the two faster speeds. Finally, during the identity condition, we found that the identity accuracy was lower than the location accuracy [ $F(1, 7) = 20.9, p < .01$ ]. Observers were sometimes able to identify the locations of the targets without being able to determine which target should go in which location.

In contrast to Pylyshyn (2004), these data show that location accuracy decreases when attempting to track a target's identity *as well as* its location. This suggests that the mechanisms supporting location tracking are not independent of identity-tracking mechanisms. Critically, this is only observed at the higher speeds. When tracking is easy, location performance is insensitive to the additional demand of tracking identities. This is presumably why Horowitz et al. (2007) found that having observers track the identities of targets did not affect their ability to track the locations of the targets as well.

## Experiment 2

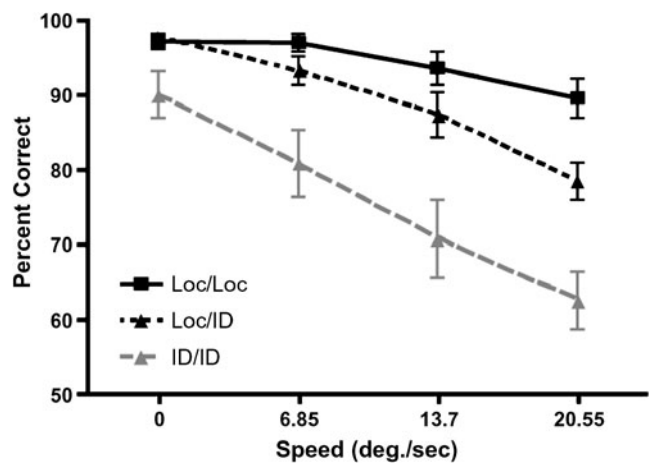
In the previous experiment, we found that asking observers to track identities as well as locations impaired location tracking only at higher speeds. In this experiment, we tried to generalize this beyond speed to other dimensions of task difficulty. Here, speed remained constant and observers tracked two, three, or four out of eight total objects.

## Method

**Participants** Eight participants were recruited from the Brigham and Women's Hospital Visual Attention Laboratory volunteer pool, as described above. The participants ranged in age from 18 to 45 years ( $M = 28.1$  years,  $SD = 12.2$ ).

**Apparatus and stimuli** The apparatus and stimuli for [Experiment 2](#) were identical to those used in [Experiment 1](#).

**Procedure** The procedure of [Experiment 2](#) was similar to the one used in [Experiment 1](#), but with two variations. First, instead of eight blocks of 40 trials, observers completed six blocks of 40 trials. As before, location- and identity-tracking conditions were run in separate blocks. The order in which these two conditions were completed was counter-balanced across observers. The second change was that the number of objects tracked varied, while the speed of the



**Fig. 2** Data from [Experiment 1](#). Accuracy is represented on the y-axis, and the speeds of the targets are represented on the x-axis. Data marked “Loc/Loc” (solid black line, squares) represent location-tracking performance (percent correct) for the location-only condition. “Loc/ID” (dashed black line, triangles) represents location-tracking performance during the identity-tracking condition. “ID/ID” (dashed gray line, triangles) represents identity performance (percent correct) during the identity-tracking condition. Error bars represent standard errors of the means

objects remained constant. Observers tracked two, three, or four of eight total objects. The tracking load remained constant within blocks and varied across blocks. In all conditions, objects moved at a rate of  $13.7^\circ/\text{s}$ , an intermediate speed in [Experiment 1](#).

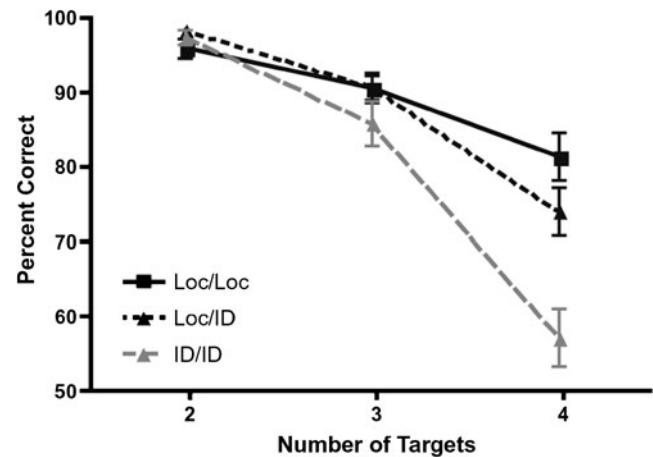
## Results and discussion

The results from [Experiment 2](#) are plotted in [Fig. 3](#). Performing a two-way within-subjects repeated measures ANOVA with Number of Targets and Condition as factors, we found an overall significant effect of target load [ $F(2, 14) = 371.9, p = .001$ ]. As in [Experiment 1](#), we found that location tracking was impaired when also maintaining identities, as compared to when observers only tracked locations [ $F(1, 7) = 12.2, p < .01$ ]. This effect again interacted with target difficulty, being larger for higher target loads [ $F(2, 14) = 4.6, p < .05$ ], such that the difference was only obtained when observers tracked four targets. Post-hoc tests showed that the difference between the conditions was only present when tracking four targets. Similar to [Experiment 1](#), during the identity condition, accuracy for tracking the identities of items was lower than accuracy for tracking the target locations [ $F(1, 7) = 58.1, p < .001$ ]. Thus, this experiment replicates the findings from [Experiment 1](#). Considered together, these two experiments confirm that tracking identities does reduce the ability to track locations, once overall task difficulty is sufficiently high.

## Experiment 3

The two previous experiments show that there is a trade-off between location tracking and identity tracking. Recall that Pylyshyn's (2004) original experiments found no difference in location tracking as a function of whether or not the observers were asked to track identities. What accounts for this discrepancy? As noted above, a key difference between our experiments and those of Pylyshyn (2004) is that object identities were continuously visible in our experiments but were visible only during the target designation phase in the Pylyshyn (2004) study. Does this factor account for the difference in results? [Experiment 3](#) was designed to answer this question by closely replicating Pylyshyn's (2004) procedures.

Instead of cartoon animals, the stimuli were cyan disks. The targets were indicated by a red square at the beginning of the trial before the motion began. Similar to Pylyshyn (2004), during this cuing phase, the targets were also labeled with the digits 1–4. When the motion sequence began, the digits disappeared.



**Fig. 3** Data from [Experiment 2](#). Accuracy is represented on the y-axis, and the speeds of the targets are represented on the x-axis. Data marked “Loc/Loc” (solid black line, squares) represent location-tracking performance (percent correct) for the location-only condition. “Loc/ID” (dashed black line, triangles) represents location-tracking performance during the identity-tracking condition. “ID/ID” (dashed gray line, triangles) represents identity performance (percent correct) during the identity-tracking condition. Error bars represent standard errors of the means

## Method

**Participants** Eight participants were recruited from the Brigham and Women’s Hospital Visual Attention Laboratory volunteer pool, as described above. The participants ranged in age from 19 to 44 years ( $M = 27.4$  years,  $SD = 7.5$ ).

**Apparatus and stimuli** The apparatus used for [Experiment 3](#) was the same as that used in the previous experiments. Instead of cartoon animals, however, the stimuli were eight cyan discs (diameter  $2.4^\circ$ ).

**Procedure** Observers completed six blocks of 40 trials. The procedure was substantially similar to the previous experiments, with the following exceptions. During the cuing phase, each target was uniquely labeled with one of the digits 1, 2, 3, or 4 in black ( $2.5^\circ$  of visual angle). In the identity-tracking condition, observers were instructed that they would have to report which digit had been presented on which target at the end of the trial. In the location-tracking condition, observers were instructed to ignore the digits. When the motion began, the numbers on the targets disappeared. Thus, during the tracking phase, all eight disks were indistinguishable from one another.

The motion algorithm was changed to be closer to Pylyshyn’s (2004) experiments. Unlike our previous experiments, objects did not bounce off each other during the motion sequence. Instead, when two objects crossed paths, they would overlap. Furthermore, the velocities of the objects varied randomly from trial to trial, ranging from

1°/s to 7°/s, with an average of 2.5°/s. Tracking duration varied randomly across trials and was either 2, 5, or 10 s.

At the end of the tracking phase, the disks stopped moving but remained visible. Responses were collected as in the previous experiments.

## Results and discussion

The results from [Experiment 3](#) are plotted in [Fig. 4](#). An ANOVA on tracking duration and condition revealed the same significant effect for trial duration [ $F(2, 14) = 22.4, p < .01$ ] as Pylyshyn (2004) observed. However, we failed to replicate Pylyshyn's initial finding, in which tracking identities did not affect location tracking. Instead, we replicated our results from [Experiments 1 and 2](#) showing that identity tracking did affect location tracking [ $F(1, 7) = 30.6, p < .001$ ]. Similar to the effects observed with speed and target load in the previous experiments, we found that the difference in location accuracy between the location and identity conditions increased with tracking duration [ $F(2, 14) = 5.9, p < .05$ ]. Post-hoc tests showed a significant difference only at the longest tracking duration. Again, accuracy for tracking the identities of items was lower than for location tracking during the identity condition [ $F(1, 7) = 79.4, p < .001$ ].

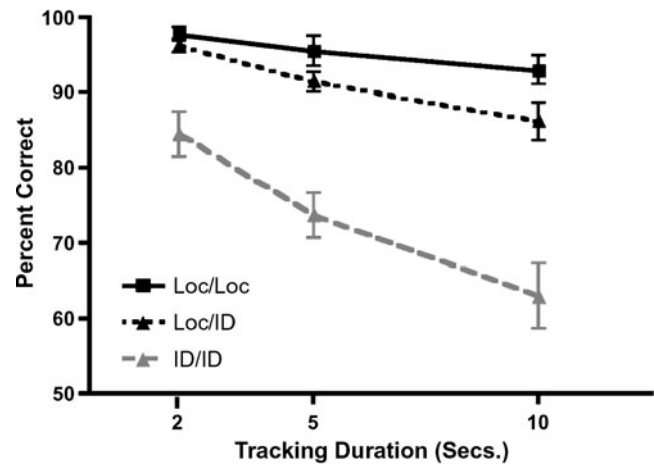
Perhaps surprisingly, the difference between our results and those obtained by Pylyshyn (2004) does not seem to be due to the continuous visibility of identities in our experiments. Instead, the difference may lie in observers' strategies. Perhaps Pylyshyn's (2004) observers had placed a high priority on the location task, only allocating resources to identity tracking when location tracking was assured? In contrast, our observers may have been allocating resources more equally between the two tasks. We test this possibility directly in [Experiment 4](#) by providing observers with explicit instructions about resource allocation.

## Experiment 4

In this experiment, observers were explicitly told to prioritize the location of the objects over the identity of the objects. However, if they felt that they could keep track of locations and allocate some resources to maintaining the identities of the targets, they were to do so. Our hypothesis was that this instruction set would produce a replication of Pylyshyn's (2004) results.

### Method

**Participants** Eight participants were recruited from the Brigham and Women's Hospital Visual Attention Laboratory



**Fig. 4** Data from [Experiment 3](#). Accuracy is represented on the y-axis, and the speeds of the targets are represented on the x-axis. Data marked “Loc/Loc” (solid black line, squares) represent location-tracking performance (percent correct) for the location-only condition. “Loc/ID” (dashed black line, triangles) represents location-tracking performance during the identity-tracking condition. “ID/ID” (dashed gray line, triangles) represents identity performance (percent correct) during the identity-tracking condition. Error bars represent standard errors of the means

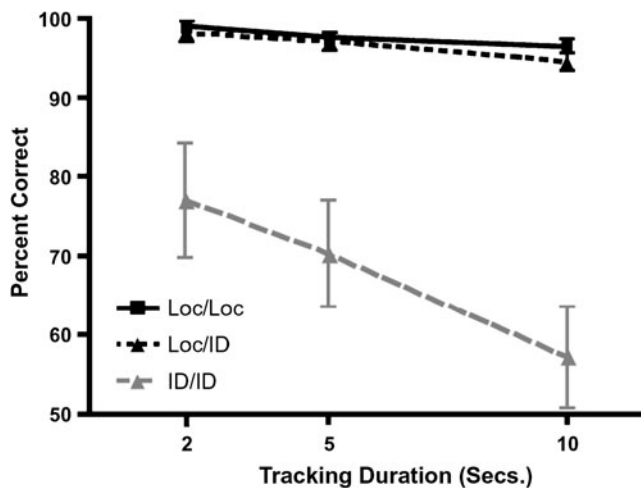
volunteer pool, as described above. The participants ranged in age from 18 to 51 years ( $M = 33.3$  years,  $SD = 9.8$ )

**Apparatus and stimuli** The apparatus and stimuli for [Experiment 4](#) were identical to those used in [Experiment 3](#).

**Procedure** The only procedural differences between [Experiment 3](#) and [Experiment 4](#) were the instructions given. In [Experiment 3](#), observers were not given any specific instructions about the strategy they should adopt. For this experiment, however, observers were explicitly told that their primary task was not to lose the locations of the targets. In the identity condition, recalling the numerical identities that went with the targets was described as the secondary task. In other words, locations were to be prioritized over identities.

## Results and discussion

The results from [Experiment 4](#) are plotted in [Fig. 5](#). With explicit instructions, we were able to replicate Pylyshyn's (2004) data. An ANOVA with Tracking Duration and Condition as factors showed that tracking performance decreased as a function of trial duration [ $F(2, 14) = 18.1, p < .001$ ] but that location tracking was not affected by identity tracking [ $F(1, 7) = 3.6, p = .1$ ]. During the identity condition, accuracy for tracking the identities of items was lower than for location tracking [ $F(1, 7) = 20.9, p < .01$ ].



**Fig. 5** Data from Experiment 4. Accuracy is represented on the y-axis, and the speeds of the targets are represented on the x-axis. Data marked “Loc/Loc” (solid black line, squares) represent location-tracking performance (percent correct) for the location-only condition. “Loc/ID” (dashed black line, solid triangle) represents location-tracking performance during the identity-tracking condition. “ID/ID” (dashed gray line, triangles) represents identity performance (percent correct) during the identity-tracking condition. Error bars represent standard errors of the means

When observers adopt the strategy of prioritizing location tracking, location performance can be maintained at the single-task level even when observers are also asked to track identities. Under this view, the successful maintenance of binding an item’s identity to its location is the result of an excess of resources above those needed for location tracking, which resources can then be allocated to identity tracking. As duration increases, the location task demands more resources, reducing identity performance.

## Experiment 5

In the previous experiment, we found that when observers were instructed to prioritize location tracking over identity tracking, location performance could be preserved while tracking identities, whereas in Experiments 1–3, with no explicit instructions, location performance tended to fall off when observers were asked to maintain target identities. This indicates that observers can strategically control the allocation of a common resource between location and identity tracking. Experiment 5 was designed to directly test this hypothesis, by asking a group of observers to perform the task under two different explicit instruction sets. For one set, observers were told to adopt the location emphasis strategy, which had them prioritize the location of targets over their identities, as in Experiment 4. For the other set, observers were told to adopt the equal-emphasis strategy,

which had them maintain the identities of the targets as well as they could during the tracking, maintaining equal priorities on locations and identities.<sup>1</sup> This is what we think observers were doing in Experiment 3, but here it was made explicit. If location and identity tracking draw upon the same resource pool, then as performance on one tracking task goes up, it should cause performance on the other to go down.

## Method

**Participants** Eight participants were recruited from the Brigham and Women’s Hospital Visual Attention Laboratory volunteer pool, as described above. The participants ranged in age from 20 to 45 years ( $M = 30.0$  years,  $SD = 9.4$ )

**Apparatus and stimuli** The apparatus and stimuli for Experiment 5 were identical to those used in Experiments 3 and 4.

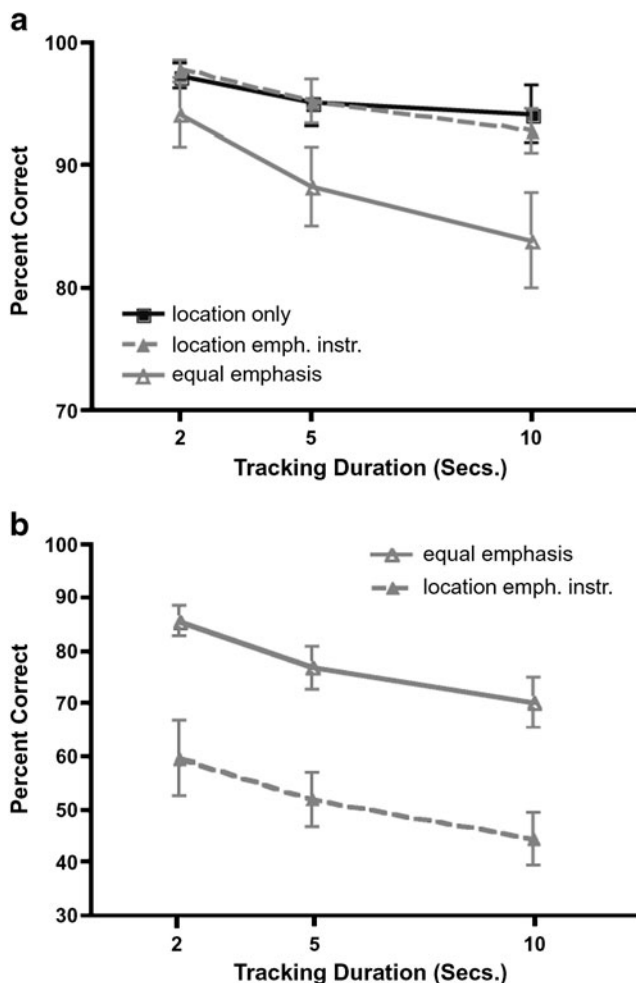
**Procedure** In terms of procedure, Experiment 5 was very similar to Experiments 3 and 4. Observers completed nine blocks of 40 trials. For three of the blocks, observers performed the location condition, while for the remaining six blocks they performed the identity condition. Critically, for three of the identity blocks, observers were told to use the equal-emphasis strategy, and for the other three blocks, they were told to use the location emphasis strategy. The order in which observers completed the experiment was counterbalanced in terms of location/identity condition and the duration of each trial. Besides the types of instructions given and the total number of blocks performed, the procedures of Experiment 5 were the same as those for Experiments 3 and 4.

## Results and discussion

The results from Experiment 5 are plotted in Fig. 6. In Fig. 6a, we see location-tracking performance as a function of tracking duration for the two different types of instructions: equal emphasis and location emphasis. Data were analyzed with an ANOVA using Tracking Duration and Instruction Condition as factors. As we found in the previous experiment, when the instruction emphasized location tracking, location-tracking accuracy did not decrease when the observers also attempted to track the identities of the targets [ $F(1, 7) = 0.04, p = .84$ ].

<sup>1</sup> Note that it was not possible to have people put greater emphasis on the identities. This is because to successfully indicate the final location of a particular target (e.g., Disk 1), observers need to bind that identity to a location—hence, equal emphasis on the location and identity tasks.

Location-tracking performance was significantly better under the location emphasis instructions than under the equal-emphasis instructions [ $F(1, 7) = 6.1, p < .05$ ]. This effect interacted with tracking duration [ $F(2, 14) = 5.1, p < .05$ ]; the difference between instruction conditions increased with tracking duration, although it was significant (by post-hoc tests) for all durations. Figure 6b shows identity-tracking performance as a function of duration. Performance declined with tracking duration [ $F(2, 14) = 20.3, p < .001$ ]. Reassuringly, observers performed better when instructed to give identity equal priority than when instructed to prioritize locations [ $F(1, 7) = 19.6, p < .01$ ]. There was no interaction between tracking duration and instructions [ $F(2, 14) = 0.02, p = .9$ ].



**Fig. 6** **a** Location-tracking performance from Experiment 5. The solid black line represents location-tracking performance in the location-only condition. The two gray lines show location-tracking performance during the identity-tracking conditions. The dashed gray line represents performance when given the location emphasis instructions, and the solid gray line represents performance for the equal-emphasis instructions. **b** Identity-tracking performance from Experiment 5. The dashed gray line represents performance when given the location emphasis instructions, and the solid gray line represents performance for the equal-emphasis instructions

These results clearly demonstrate that resources can be fluidly allocated between the location- and identity-tracking mechanisms. When observers are given the location emphasis instructions, location performance increases while identity performance decreases, relative to when they are given the equal-emphasis instructions.

## Experiment 6

Experiments 4 and 5 demonstrated that observers could flexibly allocate a common resource between the location- and identity-tracking mechanisms when the identities of the stimuli were not visible during the tracking phase. Experiment 6 was designed to see whether these results would generalize to the case when identities are continuously visible. In addition, instead of changing tracking duration, we varied the speed at which the objects moved. Thus, we employed the instruction procedures of Experiment 5 with the stimuli and displays of Experiment 1.

### Method

**Participants** Eight participants were recruited from the Brigham and Women's Hospital Visual Attention Laboratory volunteer pool, as described above. The participants ranged in age from 20 to 42 years ( $M = 26.9$  years,  $SD = 7.4$ ).

**Apparatus and stimuli** The apparatus and stimuli for Experiment 6 were identical to those used in Experiments 1 and 2.

**Procedure** Procedurally, Experiment 6 was almost identical to Experiment 5. There were three types of blocks: (1) location tracking, (2) identity tracking with equal-emphasis instructions, and (3) identity tracking with location emphasis instructions.

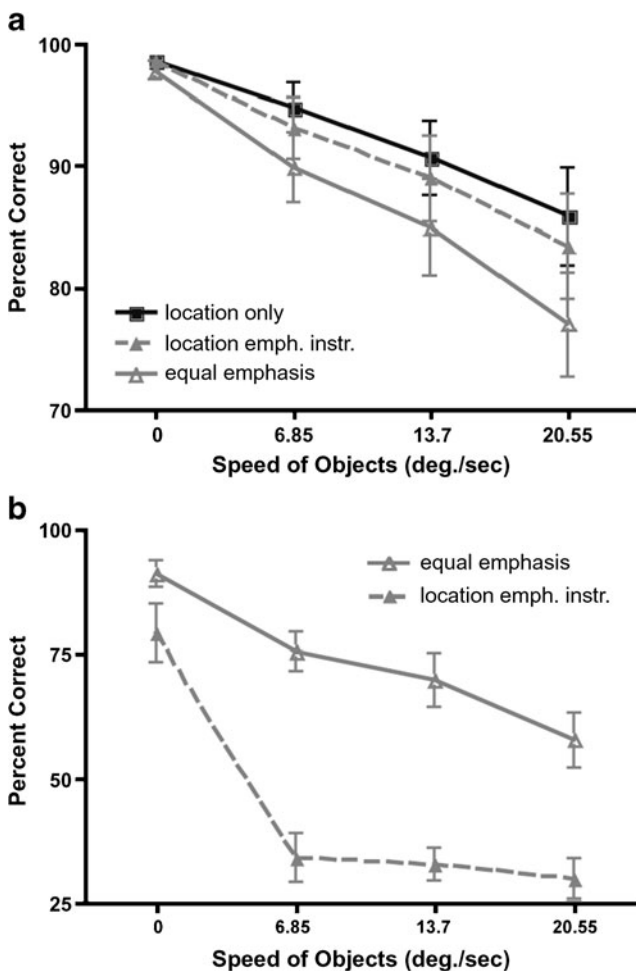
### Results and discussion

The results from Experiment 6 are plotted in Fig. 7 as a function of object speed. An ANOVA on location accuracy with Speed and Instructions as factors showed that, as seen in the previous two experiments, location tracking (Fig. 7a) did not decrease when also attempting to maintain identities when the observer instructions emphasized location tracking [ $F(1, 7) = 1.8, p = .2$ ]. As in Experiment 5, location performance differed depending on the type of instructions given [ $F(1, 7) = 10.2, p < .05$ ]. In addition, performance decreased as a function of target speed [ $F(3, 21) = 19.3, p < .01$ ]. However, instructions and target speed did not interact [ $F(3, 21) = 3.4, p = .5$ ].



The same overall pattern was observed for identity performance (Fig. 7b). There was a difference in identity accuracy depending on the instructions [ $F(1, 7) = 83.73$ ,  $p < .001$ ] and the speed of the objects [ $F(3, 21) = 39.9$ ,  $p < .001$ ]. In this case, these two effects interacted such that the instructions effect was greater as the speed of the objects increased [ $F(3, 21) = 11.3$ ,  $p < .01$ ]. Post-hoc tests showed differences between the two instruction conditions only when the objects moved, not when they remained stationary throughout the trial.

Even when object identities are continuously visible, observers can flexibly trade off resources between tracking locations and tracking identities.



**Fig. 7** **a** Location-tracking performance from Experiment 6. The solid black line represents location-tracking performance in the location-only condition. The two gray lines show location-tracking performance during the identity-tracking conditions. The dashed gray line represents performance when given the location emphasis instructions, and the solid gray line represents performance for the equal-emphasis instructions. **b** Identity-tracking performance from Experiment 6. The dashed gray line represents performance when given the location emphasis instructions, and the solid gray line represents performance for the equal-emphasis instructions

## Experiment 7

A final possibility that we have not yet considered is that the trade-off might not be due to the allocation of resources during tracking, but to the nature of the response at the end of the trial. There is an asymmetry between the responses in the location-tracking conditions, in which participants were free to click on the four targets in whatever order they chose, and the identity-tracking conditions, in which they were instructed to click on particular targets in sequence (e.g., “Click on the Zebra” followed by “Click on the Kangaroo” in Exps. 1 and 2). It is possible that this asymmetry might have interacted with our instructions to create the trade-off observed in Experiments 4–6—for example, due to a deterioration of identity–location bindings in memory. Note that Pylyshyn (2004) used a free-recall method in which participants would press a key to say which item they were going to select and then use the mouse to indicate the location of the target. For this experiment, we used the same response method as Pylyshyn (2004) to ensure that maintaining identity is the source of task interference.

### Method

**Participants** Sixteen participants were recruited from the Brigham and Women’s Hospital Visual Attention Laboratory volunteer pool, as described above. The participants ranged in age from 18 to 42 years ( $M = 25.4$  years,  $SD = 6.3$ ).

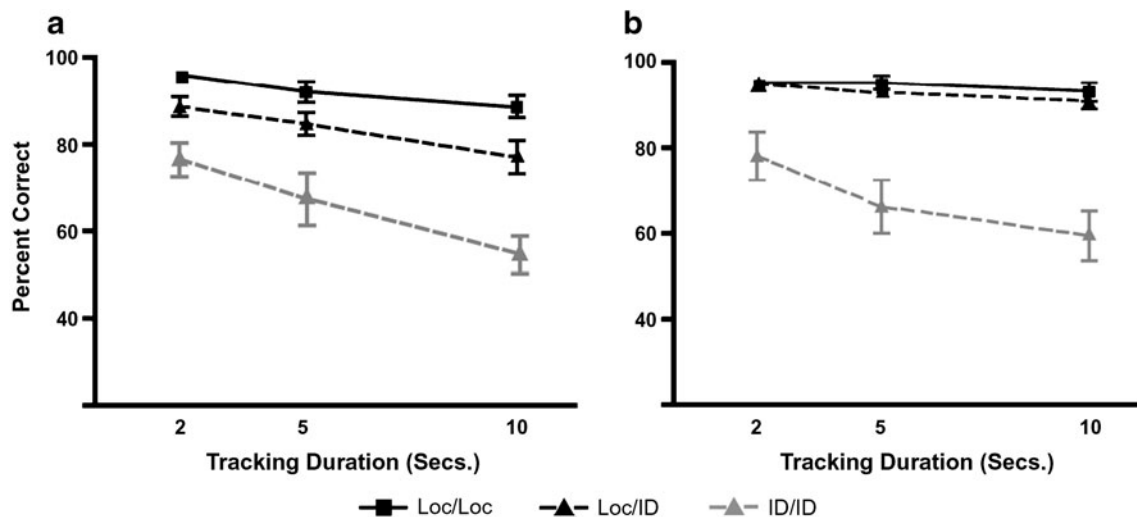
**Apparatus and stimuli** The apparatus and stimuli for Experiment 7 were identical to those used in Experiments 3, 4, and 5.

**Procedure** Procedurally, Experiment 7 was basically a combination of Experiments 3 and 4. One group of 8 participants completed the two conditions of tracking (i.e., location and identity tracking) without any special instructions. The other of 8 participants were given the location-biased instructions that were used in Experiments 4–6.

For both groups of participants, the manner in which they provided their responses was tailored to match those used in Pylyshyn’s (2004) study. At the beginning of each trial, four disks were identified as targets and had the numbers 1, 2, 3, and 4 in them. At the end of the trial, participants would click on a particular disk with the mouse and then press a key on the keypad to identify that object’s identity. This was repeated four times, once for each unique target.

### Results and discussion

The results from Experiment 7 are plotted in Fig. 8 as a function of tracking duration. The left panel shows the data



**Fig. 8** (Left) Performance in Experiment 7 for participants who were given no explicit instructions. Accuracy is represented on the  $y$ -axis, and the tracking duration is represented on the  $x$ -axis. (Right) Performance in Experiment 7 for participants who were given instructions that emphasized target location during the identity condition. Data marked “Loc/Loc” (solid black lines, squares)

represent location-tracking performance (percent correct) for the location-only condition. “Loc/ID” (dashed black lines, triangles) represents location-tracking performance during the identity-tracking condition. “ID/ID” (dashed gray lines, triangles) represents identity performance (percent correct) during the identity-tracking condition. Error bars represent standard errors of the means

for participants given no specific emphasis instructions. An ANOVA on location accuracy with Condition and Tracking Duration as factors yielded a significant effect for trial duration [ $F(1, 7) = 32.7, p < .001$ ]. Furthermore, location tracking was significantly worse when having to also track identities than when tracking locations only [ $F(1, 7) = 55.7, p < .001$ ]. This effect interacted with trial duration [ $F(1, 7) = 14.5, p < .05$ ], indicating that the difference (while significant for all points by post-hoc tests) increased with tracking duration. During the identity condition, accuracy was lower for tracking the identities of items than for location tracking [ $F(1, 7) = 92.9, p < .001$ ]. The right panel of Fig. 8 shows data for the participants who were explicitly told to prioritize the locations of the items during the identity-tracking condition. In this case, there is also an effect for trial duration [ $F(1, 7) = 14.7, p < .01$ ]. However, location tracking was not affected by identity tracking [ $F(1, 7) = 1.8, p = .22$ ]. During the identity condition, accuracy for tracking the identities of items was lower than for location tracking [ $F(1, 7) = 21.6, p < .01$ ].

The results of Experiment 7 clearly demonstrate that the differences between our results and Pylyshyn’s (2004) results were not due to the manner in which participants gave their responses. Indeed, both our response method and Pylyshyn’s yielded the same pattern of results.

## General discussion

In this article, we addressed the question of whether identity and location tracking are carried out by two independent mechanisms or whether the two tasks share a common resource. We did this using several variations on the MOT and MIT paradigms. Object identities could be visible or invisible during the tracking phase, while difficulty was manipulated via speed, tracking duration, and target load. When identities were tracked, we computed two measures of accuracy: *location* and *identity* accuracy. This allowed us to investigate what effect, if any, identity tracking has on location tracking.

Experiments 1–3 showed that location accuracy decreases when observers also attempt to maintain target identities. In other words, identity tracking impairs location tracking. Experiments 4–6 showed that observers can voluntarily distribute their mental resources across the location- and identity-tracking tasks. In sum, these results show that the mechanisms underlying location tracking and identity tracking share a common resource pool. This result is consistent with recent studies of MOT and visual working memory that have shown that mental resources can be distributed flexibly, depending on task demands (Alvarez & Franconeri, 2007; Bays & Husain, 2008; Iordanescu, Grabowecky, & Suzuki, 2009). However, these earlier studies focused exclusively on

the allocation of resources *across* targets, whereas our results show that such allocation can also be seen *within* targets (i.e., allocating resources to a particular target's location or identity, depending on the instructions).

A number of studies have documented that observers have some access to the identities (Horowitz et al., 2007; Oksama & Hyönä, 2004, 2008) or features (Howard & Holcombe, 2008) of tracked objects. The novel contribution of this study is that we demonstrated a trade-off between location and identity tracking. The two previous studies that looked for such a trade-off failed to observe one (Horowitz et al., 2007; Pylyshyn, 2004). We argue that Pylyshyn (2004) did not observe a trade-off because his observers prioritized location tracking, whereas Horowitz et al. did not observe a trade-off because they did not make the task sufficiently challenging. In all of our experiments, the size of the trade-off increased with task difficulty, whether defined in terms of tracking load, speed, or duration.

As we have noted, working memory is tightly linked to tracking (Allen, et al., 2006; Fougne & Marois, 2006; Howe et al., 2009; Oksama & Hyönä, 2004). Thus, one way to accommodate our findings would be to propose that target positions and identities are maintained in a common visual short-term memory (VSTM) store. While VSTM is often thought to comprise a fixed set of “slots” that hold holistic objects, irrespective of content (Awh, Barton, & Vogel, 2007; Cowan, 2001; Irwin, 1992; Luck & Vogel, 1997), some theorists have suggested that VSTM might instead be limited by resolution (Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005), such that the more complex the objects, the fewer that can be stored.

Zhang and Luck (2008) recently proposed two hybrid models of visual working memory that combine aspects of both approaches. The *slots + resources* model claims that there are a fixed number of slots in memory, but a fluid resource can be distributed anisotropically between these slots. The *slots + averaging* model states that when the number of items to be remembered is less than the number of available slots, multiple slots can be assigned to one item. Assigning multiple slots to an item increases resolution, because each individual slot provides independent information about the target object. These hybrid models of working memory can be easily applied to our MIT task. However, although Zhang and Luck decided in favor of the *slots + averaging* model, on the basis of the distributions of color errors in a VSTM task, our data seem more compatible with the *slots + resources* model. While the *slots + averaging* model handles redistributing accuracy between objects well, it is more difficult to see how it could explain redistributing accuracy between the attributes of a

single object, since the representational capacity of each individual slot is claimed to be fixed. In other words, as the *slots + averaging* model currently stands, when a slot is dedicated to the representation of an object, it is by definition representing the *whole* object. Assigning multiple slots to the same object would therefore improve the precision of all of that object's attributes.

One problem with the *slots + resources* account is that it runs counter to the evidence that spatial and object information are stored in separate neural buffers (Klauer & Zhao, 2004; Postma & De Haan, 1996; Ruchkin, Johnson, Grafman, Canoune, & Ritter, 1997; Vicari, Bellucci, & Carlesimo, 2003), not to mention the well-established distinction between “what” and “where” cortical processing streams (Goodale & Milner, 1992). Instead of a common *representation*, position and identity information might rely on a common *process*. The key difference between tracking and memory tasks is that tracking requires continual (or at least frequent) updating of the representation. The updating mechanism may be a more centralized resource that feeds in to both position and identity stores (Klauer & Zhao, 2004; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Vecchi, Monticellai, & Cornoldi, 1995). Studies of individual differences (e.g., Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005) might be a useful approach to disentangle these alternatives. It is not clear at the present time whether the shared resource described here is purely visual or is more task-general. Previous work has shown that MOT can interfere with other visual tasks, such as working memory (Fougne & Marois, 2006, 2009), but it can also affect an auditory discrimination task (Tomblu & Seiffert, 2008). Future research is needed to more specifically identify the scope and limits of the resources involved in all of these tasks and their relationships to one another.

## Conclusion

Previous studies have claimed that the locations of moving objects are tracked independently of their identities. Here, we presented multiple experiments that investigated the cost of attempting to simultaneously track objects' locations and identities. Contrary to previous findings (Horowitz et al., 2007; Pylyshyn, 2004), we found that observers are worse at tracking target locations when they are also required to track the target identities. Moreover, observers can trade off performance in the two tasks. This suggests that identity tracking and location tracking draw upon a common resource.

## Appendix

### Stimuli Used in Experiments 1, 2, and 6



## References

- Allen, R., McGeorge, P., Pearson, D. G., & Milner, A. D. (2006). Multiple-target tracking: A role for working memory? *The Quarterly Journal of Experimental Psychology*, *59*, 1101–1116.
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106–111.
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(13), 1–10. 14.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*, 622–628.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*, 349–354.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*, 87–185.

- Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*, 1127–1133.
- Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory. *Psychological Science*, *17*, 526–534.
- Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition*, *17*, 48–66.
- Franconeri, S. L., Jonathan, S., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by object spacing, not speed, time, or capacity. *Psychological Science*, *21*, 920–925.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 20–25.
- 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.
- Howard, C., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research*, *48*, 1164–1180.
- Howe, P. D., Horowitz, T. S., Morocz, I. A., Wolfe, J., & Livingstone, M. S. (2009). Using fMRI to distinguish components of the multiple object tracking task. *Journal of Vision*, *9*(4), 1–11. 10.
- Imaruoka, T., Saiki, J., & Miyauchi, S. (2005). Maintaining coherence of dynamic objects requires coordination of neural systems extended from anterior frontal to posterior parietal brain cortices. *Neuroimage*, *26*, 277–284.
- Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*, *9*(4), 1–12. 1.
- Irwin, D. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 307–317.
- Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General*, *133*, 355–381.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Makovski, T., & Jiang, Y. (2009). Feature binding in attentive tracking of distinct objects. *Visual Cognition*, *17*, 180–194.
- Makovski, T., Vazquez, G. A., & Jiang, Y. V. (2008). Visual learning in multiple-object tracking. *PLoS ONE*, *3*, e2228.
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.
- Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*, 631–671.
- Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology*, *56*, 237–283.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Postma, A., & De Haan, E. (1996). What was where? Memory for object locations. *The Quarterly Journal of Experimental Psychology*, *49*, 178–199.
- Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*, 65–97.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*, 127–158.
- Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Visual Cognition*, *11*, 801–822.
- Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. Cambridge: MIT Press.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197.
- Ruchkin, D. S., Johnson, R., Grafman, J., Canoune, H., & Ritter, W. (1997). Multiple visuospatial working memory buffers: Evidence from spatiotemporal patterns of brain activity. *Neuropsychologia*, *35*, 195–209.
- Saiki, J. (2002). Multiple-object permanence tracking: Limitation in maintenance and transformation of perceptual objects. *Progress in Brain Research*, *140*, 133–148.
- 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: MIT Press.
- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.
- Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple-object tracking. *Cognition*, *108*, 1–25.
- Vecchi, T., Monticellai, M. L., & Cornoldi, C. (1995). Visuo-spatial working memory: Structures and variables affecting a capacity measure. *Neuropsychologia*, *33*, 1549–1564.
- Vicari, S., Bellucci, S., & Carlesimo, G. A. (2003). Visual and spatial working memory dissociation: Evidence from Williams syndrome. *Developmental Medicine and Child Neurology*, *45*, 269–273.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748–751.
- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*, 500–503.
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48–64.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.
- Zhang, W., & Luck, S. J. (2008). Discrete, fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235.