



UvA-DARE (Digital Academic Repository)

The more often you see an object, the easier it becomes to track it

Pinto, Y.; Howe, P.D.L.; Cohen, M.A.; Horowitz, T.S.

DOI

[10.1167/10.10.4](https://doi.org/10.1167/10.10.4)

Publication date

2010

Document Version

Final published version

Published in

Journal of Vision

[Link to publication](#)

Citation for published version (APA):

Pinto, Y., Howe, P. D. L., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of Vision*, 10(10), 4. <https://doi.org/10.1167/10.10.4>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

The more often you see an object, the easier it becomes to track it

Yair Pinto

Brigham and Women's Hospital, Boston, MA, USA, &
Harvard Medical School, Boston, MA, USA



Piers D. L. Howe

University of Melbourne, Parkville, Victoria, Australia



Michael A. Cohen

Harvard University, Cambridge, MA, USA



Todd S. Horowitz

Brigham and Women's Hospital, Boston, MA, USA, &
Harvard Medical School, Boston, MA, USA



Is it easier to track objects that you have seen repeatedly? We compared repeated blocks, where identities were the same from trial to trial, to unrepeated blocks, where identities varied. People were better in tracking objects that they saw repeatedly. We tested four hypotheses to explain this repetition benefit. First, perhaps the repeated condition benefits from consistent mapping of identities to target and distractor roles. However, the repetition benefit persisted even when both the repeated and the unrepeated conditions used consistent mapping. Second, repetition might improve the ability to recover targets that have been lost, or swapped with distractors. However, we observed a larger repetition benefit for color–color conjunctions, which do not benefit from such error recovery processes, than for unique features, which do. Furthermore, a repetition benefit was observed even in the absence of distractors. Third, perhaps repetition frees up resources by reducing memory load. However, increasing memory load by masking identities during the motion phase reduced the repetition benefit. The fourth hypothesis is that repetition facilitates identity tracking, which in turn improves location tracking. This hypothesis is consistent with all our results. Thus, our data suggest that identity and location tracking share a common resource.

Keywords: multiple object tracking, identity, binding, repetition, familiarity

Citation: Pinto, Y., Howe, P. D. L., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of Vision*, 10(10):4, 1–15, <http://www.journalofvision.org/content/10/10/4>, doi:10.1167/10.10.4.

Introduction

There are many real-world tasks that require us to keep track of moving objects. These range from specialized tasks such as air traffic control and radar operation (Allen, McGeorge, Pearson, & Milne, 2004) to everyday tasks such as driving in traffic and monitoring your children at the beach.

Our research question concerns the observer's experience with the identity of the tracked objects. Specifically, is it easier to track objects that you have seen repeatedly? Imagine you place your pet fish in an aquarium with your neighbor's pet fish. Is it easier to monitor your own pet fish than your neighbor's? The answer to this question holds straightforward practical implications (for starters, perhaps you should not let others monitor your pet fish when it really matters), but it also elucidates the mechanisms that drive monitoring moving objects.

The laboratory method for this sort of tracking is the multiple object tracking paradigm (MOT, Pylyshyn & Storm, 1988). In a typical MOT trial, the observer is presented with an array of identical objects (e.g., gray disks).

The targets are briefly highlighted, after which all objects move independently for several seconds. The observer must then indicate the location of the targets (see Figure 1). During the moving phase, all the objects are identical, so the observer can do the task only by continuously attending to the targets. In most cases, observers can track 3–5 targets (Cowan, 2001; Intriligator & Cavanagh, 2001; Yantis, 1992), though tracking performance can be modulated by the speed of the objects and how close they come to one another (Alvarez & Franconeri, 2007; Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008).

While the standard MOT paradigm has taught us a great deal about object tracking (for reviews, see Cavanagh & Alvarez, 2005; Scholl, 2009), it is not naturalistic, in the sense that in the real world tracked objects are often distinct rather than identical. The first investigation of MOT where the targets had distinct identities was performed by Pylyshyn (2004). The paradigm was similar to that described above except that, at the start of each trial, each target was labeled with a digit. The digits then disappeared and were not visible during the motion phase. At the end of the trial, the observer was asked to indicate first which items were the targets and then which digit

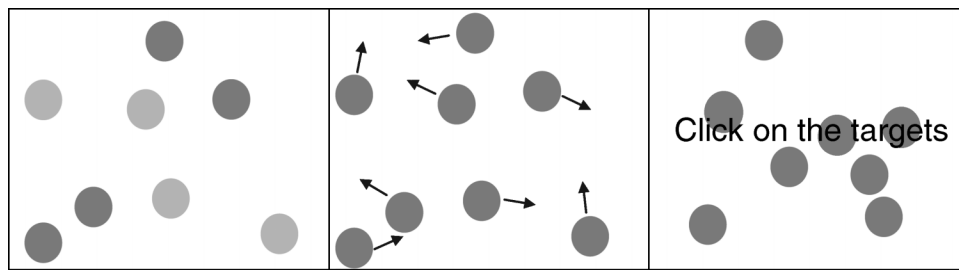


Figure 1. An example of a typical MOT trial. Four disks are highlighted by turning light gray. After a little while, all disks turn dark gray and start moving. At the end of the trial, the participant is asked to click on the target disks.

went with which target. Pylyshyn found that the capacity for target identities was substantially lower than the capacity for target locations. These data suggest that there is a *binding problem* (Wolfe & Cave, 1999) for target identities in MOT (see also Imaruoka, Saiki, & Miyauchi, 2005; Saiki, 2002, 2003).

One could argue that Pylyshyn's (2004) paradigm is still unrealistic because in the real world target identities are usually continuously visible. Oksama and Hyönä (2004) addressed this issue in their *multiple identity tracking* (MIT) experiments in which the stimuli were distinct objects (or pseudo-objects) that were continuously visible. At the end of the trial, all the objects were masked, and the observer was asked to report the identity of a randomly selected target. They found that observers could track approximately four objects without confusing their identities, much more than the capacity reported by Pylyshyn (2004), suggesting that the identity–location binding problem is at least partially ameliorated by making the targets continuously visible during the moving phase.

If continuous visibility completely eliminated the binding problem, then we would expect tracking capacity for MIT to be the same as that for MOT. However, Horowitz et al. (2007) reported that this was not the case. In their experiments, the stimuli were cartoon animals. In the *specific* condition, observers were asked to identify the location of particular objects, similar to Oksama and Hyönä's (2004) setup. In the *standard* condition, observers were asked to identify the locations of all the targets, without specifying which target was where. Capacity in the specific condition ranged from 1.4 to 2.6 objects, whereas in the standard condition it ranged from 2.3 to 3.4 objects. Thus, even when the object identities were continuously visible during the movement phase, the tracking capacity for MIT was still less than that for MOT.

As MOT studies move away from identical stimuli toward more ecologically valid studies that utilize distinct objects, we need to consider not only the influence of stimulus factors on tracking accuracy (e.g., velocity, Alvarez & Franconeri, 2007; spacing, Franconeri et al., 2008; color and shape, Makovski & Jiang, 2009) but also the interplay between tracking and memory (Allen,

McGeorge, Pearson, & Milne, 2006; Fougne & Marois, 2006; Oksama & Hyönä, 2004; Saiki, 2003).

Oksama and Hyönä's (2008) study is a clear example of the importance of non-stimulus factors driving tracking performance. Their data suggest that it is easier to track objects that you are more familiar with. They found tracking performance to be significantly better for real, common objects (drawings from Snodgrass & Vanderwart, 1980) than for pseudo-objects (drawings of object-like stimuli, from Kroll & Potter, 1984). This finding directly relates to our research question since it suggests that it is easier to track objects that you have had more exposure to. However, their results were not conclusive. First, in their design, familiarity is confounded by nameability. Nameable, real objects might have been easier to remember, thereby reducing memory load and consequently improving tracking performance. Second, there might be low-level visual differences between the real and the pseudo-objects Oksama and Hyönä used. These differences could have contributed to the observed improvement in tracking performance.

We avoided these issues by using the same set of stimuli in all conditions, thus equating for nameability and perceptual effects. We then compared performance under repeated (same target and distractors throughout a block) and unrepeated (targets and distractors randomly sampled throughout a block) conditions. We found that repeating item identities significantly improved tracking performance. We considered four possible explanations. First, identity repetition might reduce target–distractor interference, since distractor identities never become target identities and vice versa. Second, identity repetition might make it easier to recover lost targets. Third, identity repetition might reduce working memory load, freeing up resources for tracking. Finally, perhaps resources can flexibly be distributed between identity and location tracking. It could be that identity repetition makes identity tracking more efficient, thus freeing more resources for location tracking.

Identity tracking here is really comprised of two processes: recognizing and maintaining the *representation* of an identity; and *binding* this representation to a position or spatial index. Thus, the visual system must maintain a

visual representation of the zebra as it moves and bind the zebra image to a particular spatial locus. It is important to think about this *visual* representation of the zebra as something different from simply *knowing* that one is tracking a zebra, a distinction that will become important in [Experiment 5](#).

Our results are most consistent with the hypothesis that repetition improves identity tracking. Whether it does so through facilitating the visual representations of targets or their binding to positions, or both, cannot be determined from our experiments. Furthermore, while we show that each of the other hypotheses is insufficient to explain the repetition benefit, neither can we exclude the possibility that they may contribute to the effect. In the [General discussion](#) section, we discuss the implication of these findings for models of MIT.

Experiment 1: Identity repetition improves tracking performance

As noted above, Oksama and Hyönä's data suggest that familiarity with the stimuli enhances tracking performance, but their result is potentially confounded with nameability and the perceptual qualities of the stimuli. The goal of our first experiment was to demonstrate that merely repeating target identities improves tracking performance.

We compared performance when all stimuli changed from trial to trial within a block (the *unrepeated* condition), to a situation where the targets and distractors remained the same on each trial throughout the block (the *repeated* condition). Crucially, each object had an equal chance of being assigned to the repeated set for all observers. Thus, there should be no visual or semantic differences between the repeated and unrepeated objects.

Methods

Participants

Eight naive, paid volunteers, ranging in age from 18 to 50 years (average 27.5 years) participated in this experiment. All participants had normal or corrected-to-normal vision.

Apparatus and stimuli

Stimuli were presented on a 21-inch monitor set to a resolution of 1024 by 768 at a refresh rate of 75 Hz, controlled by a Macintosh G5 computer running Mac OS 10.4. The experiment was programmed in Matlab 7.5 (The MathWorks) using the Psychophysics Toolbox routines (Brainard, 1997; Pelli, 1997). Participants were seated approximately 57.4 cm from the monitor; at this distance, 1 cm on the screen subtends 1 degree of visual angle ($^{\circ}$). The stimulus field consisted of eight cartoon animals ($2.37^{\circ} \times 2.37^{\circ}$). Animals moved at a speed of 17.8°/s in straight lines, except when they bounced off each other or the sides of an imaginary window ($32.7^{\circ} \times 22.5^{\circ}$). At the end of a trial, each animal was masked by a red (luminance: 10.5 cd/m², as measured with a Tektronix photometer, and CIE (x, y) coordinates of (0.617, 0.349)) $2.37^{\circ} \times 2.37^{\circ}$ square. The background was white (0.289, 0.317, luminance: 63.2 cd/m²).

Procedure

The experiment consisted of 8 blocks of 30 trials. On all trials, participants tracked four targets among four distractors. There was a set of 22 cartoon animals (see [Figure A1](#)). In four of the blocks, targets and distractors randomly varied from trial to trial (the “unrepeated” condition); in the other four blocks, both targets and distractors remained the same on every trial (the “repeated” condition). Half of the participants started with the repeated condition, the other half started with the unrepeated condition. The basic trial procedure is illustrated in [Figure 2](#). Each trial started with highlighting the targets in red boxes ($3.57^{\circ} \times 3.57^{\circ}$, luminance: 10.5 cd/m², (0.617, 0.349)) for 3 s. All items then moved for a randomly determined time between 2 and 12 s. The length of the trial was unpredictable in order to ensure that participants were continually tracking rather than adopting a strategy of locating the targets just before the end of the trial. All items were then masked with red squares. Participants were asked to click on each specific target in turn (e.g., “where is the zebra?,” followed by “where is the turtle?,” etc.). Participants were instructed to take their time and respond as accurately as possible. The experiment was performed without breaks and took approximately 90 min.

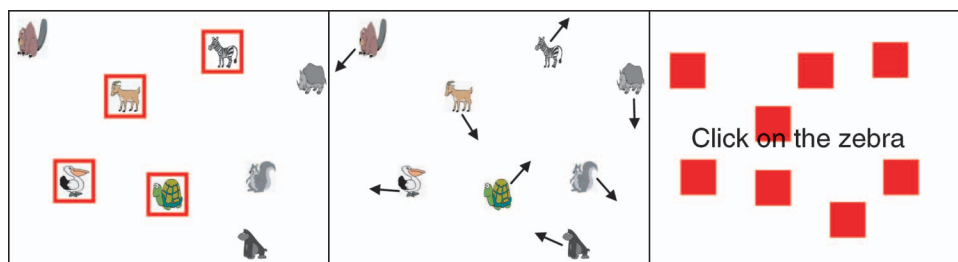


Figure 2. An example of the task we used in [Experiment 1](#). The trial started by highlighting the targets. During motion, all identities would remain visible. The trial stopped at an unpredictable moment, at which point participants were asked to click on specific targets.

Data analysis

We analyzed two dependent measures: *location* accuracy and *identity* accuracy. Location accuracy refers to participants' ability to distinguish target from distractor locations, whereas identity accuracy refers to their ability to know which target is where. Imagine that a participant is asked to track the alligator, the camel, the fox, and the zebra, among distractors that include the turtle, the elephant, the horse, and the tiger. If the participant was asked to click on, say, the zebra, and she clicked on the camel, this would be considered a hit with respect to location accuracy, because it indicates that she knew that the item was a target, but as a miss with respect to identity accuracy, because she did not know which target it is. Note that with this definition, identity accuracy can never exceed location accuracy.

Results and discussion

Location and identity accuracies are plotted in [Figure 3](#) as a function of repetition. Performing a 2-way within-subjects ANOVA, we found that location accuracy was significantly greater than identity accuracy ($F(1, 7) = 22.46$, $MSE = 75.48$, $p < 0.005$, $\eta^2 = 0.762$), and accuracy for repeated objects was greater than accuracy for unrepeated objects ($F(1, 7) = 17.52$, $MSE = 56.72$, $p < 0.005$, $\eta^2 = 0.715$). Furthermore, these two effects interacted, such that the effect of repetition was greater for identity accuracy than for location accuracy ($F(1, 7) = 16.53$, $MSE = 12.10$, $p = 0.005$, $\eta^2 = 0.702$). Finally, post-hoc tests showed that both identity accuracy and location accuracy were significantly higher in the repeated case ($F(1, 7) = 22.31$, $p < 0.005$, $\eta^2 = 0.761$ and $F(1, 7) = 6.85$, $p < 0.05$, $\eta^2 = 0.494$, respectively).

The main result of [Experiment 1](#) is that identity accuracy increased from roughly 64% when tracking unrepeated targets to roughly 81% when tracking repeated targets.

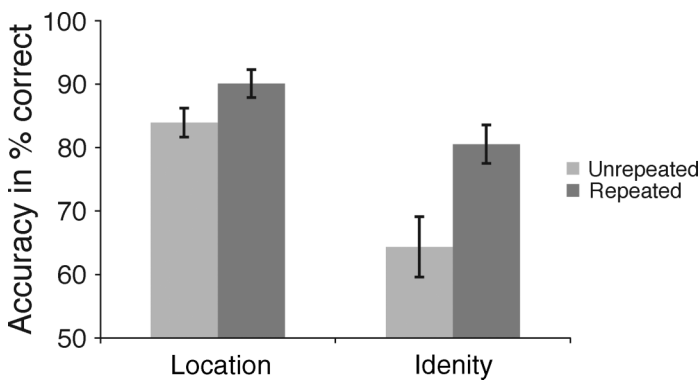


Figure 3. Location and identity accuracy data from [Experiment 1](#) in the unrepeated (light gray bars) and the repeated (dark gray bars) conditions. Error bars (in this and subsequent figures) indicate the standard error of mean.

This indicates that it is indeed easier to track unique targets if you have seen them more often (an effect that we will refer to as the repetition benefit). Another noteworthy result is that location accuracy also increased for repeated items. Thus, participants were not only better at tracking identities but also at tracking locations.

Experiment 2: Consistent mapping does not eliminate the repetition effect

What is the source of the repetition benefit demonstrated in [Experiment 1](#)? One possibility is that this benefit has nothing to do with repetition per se, but more with consistency of target and distractor roles. In the unrepeated condition in our experiments, any given item could serve as both a distractor and a target, on different trials, while in the repeated condition an item's role was always fixed. This aspect of the design could potentially create a repetition benefit by itself in a number of ways. For example, visual search studies have shown that inconsistent mapping (where targets and distractors can swap roles) leads to impaired performance relative to consistent mapping (Schneider & Shiffrin, 1977). Other visual search studies have demonstrated that a distractor is more distracting if it served as a target on the previous trial (target-distractor priming, Pinto, Olivers, & Theeuwes, 2005). Conversely, negative priming studies (Tipper, 1985; Tipper, Weaver, & Wright, 1998) have also demonstrated that ignored distractors are harder to respond to when they show up as targets on the next trial. Thus, it is possible that the repetition benefit has nothing to do with identity repetition per se but is simply a consequence of a design flaw in our experiments.

In [Experiment 2](#), we addressed this concern by introducing two new conditions. In both conditions, target and distractor identities were chosen from separate pools, ensuring that targets never served as distractors and vice versa. If the repetition benefit is really a benefit for consistent mapping, then we would expect that tracking performance in these new conditions would be as good as tracking in the repeated condition. However, if the repetition benefit is indeed due to the repetition of the target identity, then we still expect tracking performance to be the best in the repeated condition.

Methods

Participants

Eight participants took part as paid volunteers. Their ages ranged from 18 to 50 (average 31.6) years, and all had normal or corrected-to-normal vision.

Stimuli and procedure

The stimuli employed in this experiment were different from the previous experiment. The stimuli were 168 pictures from the MIT objects database containing 2400 pictures of objects (<http://cvcl.mit.edu/MM/download.html>).

There were four conditions: *repeated*, *standard unrepeated*, *unrepeated Consistent Mapping 1 (CM1)*, and *unrepeated Consistent Mapping 2 (CM2)*. The repeated condition was similar to that of [Experiment 1](#), where throughout the block targets and distractors had fixed identities (the 8 identities of targets and distractors were randomly selected from the pool of 168 pictures). Similarly, the standard unrepeated condition was designed to mimic the unrepeated condition from the previous experiment. We first randomly selected 40 of the 160 remaining pictures. Then, on each trial the four target and four distractor identities were randomly chosen from this pool of 40 pictures. In the CM1 condition, we randomly selected 80 of the remaining 120 pictures. Forty pictures were then randomly assigned to the target pool and the remaining 40 pictures to the distractor pool. On each trial, the four target identities were randomly drawn from the target pool, whereas the four-distractor identities were randomly drawn from the distractor pool. This ensured that the targets could never serve as distractor and vice versa. Furthermore, the number of target repetitions in this condition was the same as the number of target repetitions in the standard unrepeated condition. Finally, in the CM2 condition the remaining group of 40 pictures was randomly divided in a group of 20 target pictures and 20 distractor pictures. Other than the number of pictures in each pool being smaller than in the CM1 condition, the CM2 and CM1 conditions were the same. Thus, the CM1 condition matched the standard unrepeated condition in terms of number of target repetitions, but used twice as many stimuli, while the CM2 condition matched the standard unrepeated condition in terms of number of stimuli, but with twice the number of target repetitions. The distribution of pictures over the different pools was done independently for each participant.

The experiment consisted of 4 blocks of 40 trials in a Latin Square design. Two participants were run on each of the four possible orders. The experiment was performed without breaks and took approximately 60 min.

Results and discussion

The data are presented in [Figure 4](#). We analyzed identity and location accuracies separately. In both cases, we compared the unrepeated and CM conditions to the repeated condition. With regard to location accuracy, the trend was toward accuracy being the highest in the repeated condition, and the lowest in the unrepeated condition, but this did not reach significance for any of the comparisons. With regard to identity accuracy, a one-way ANOVA with the factor condition (repeated, standard unrepeated, CM1,

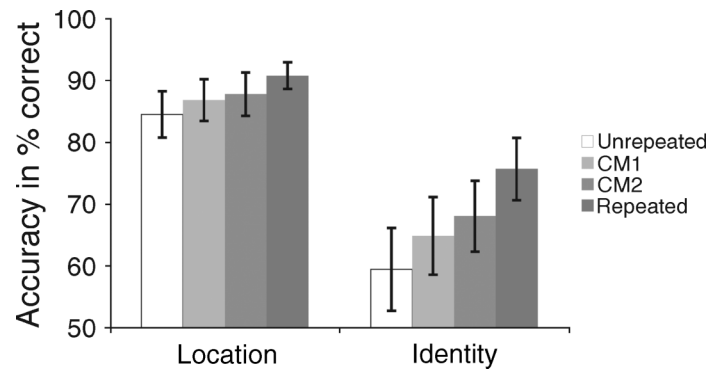


Figure 4. Location and identity accuracy data from [Experiment 2](#). From left to right, the standard unrepeated, CM1, CM2, and repeated conditions are depicted.

and CM2) revealed a main effect ($F(3, 21) = 12.83$, $MSE = 28.774$, $p < 0.0001$, $\eta^2 = 0.647$). A within-subjects contrast, contrasting all conditions to the repeated condition, revealed that accuracy in the repeated condition was higher than in all other conditions (for all conditions $F > 7.5$, all $p < 0.05$, all $\eta^2 > 0.5$).

Thus, the effects of consistent mapping do not account for the entire repetition benefit. Identity accuracy was significantly higher in the repeated condition than in both of the consistent mapping conditions. Even when target and distractor identities were strictly separated, so that inconsistent mapping, target–distractor priming, and negative priming were excluded, tracking performance was still better in the repeated than in the unrepeated condition. This shows that the repetition benefit is, at least partly, due to repeating target identities.

Another possible explanation is that the repetition benefit has the same cause as the uniqueness benefit. Previous studies (Horowitz et al., 2007; Makovski & Jiang, 2009) have shown that it is easier to track unique objects, relative to identical objects. Both studies suggested that this is because if the observer lost track of a target, he would know that he had lost, for example, the zebra, and could then search for a zebra-like item in the field to recover that target. Perhaps repeated targets are easier to track because repetition improves observers' memory for the target set, making it easier both to recognize which target is missing and to search for the missing target.

Alternatively, the repetition benefit might reflect a shift in resources from identity tracking to position tracking. If we posit that the location and identity tracking aspects of the task draw on a common resource, then anything that makes identity tracking easier would free up resources for improved position tracking. For instance, if the same identity is repeatedly bound to the same position, or object file (Kahneman, Treisman, & Gibbs, 1992; Pylyshyn, 2001), this operation may become easier over time, allowing the observer to concentrate more resources on position tracking.

Contrasting these two hypotheses required a change from the cartoon animal and real object stimuli to a more controlled stimulus set, following Makovski and Jiang (2009).

Experiment 3: Repetition effects on tracking of conjunctions and features

The purpose of Experiment 3 was twofold. First, we wanted to determine whether the repetition benefit would replicate with simple stimuli that presumably result in a lower binding load. Second, we wanted to test whether the repetition benefit could be entirely explained by improved search for lost targets. We therefore designed a condition in which each stimulus was defined by a unique feature, and a condition in which each stimulus was defined by a conjunction of two features. In the feature condition, each stimulus was a square of a unique color: red, blue, yellow, green, black, white, orange, or purple. In the conjunction condition, stimuli were squares divided vertically such that the left and right halves were of different colors. Each stimulus was a unique combination of two colors.

In the feature condition the binding load is lower than in the conjunction condition (Treisman, 2006), so if the repetition benefit is based on improved binding, the repetition benefit should be substantially lower in the feature condition than in the conjunction condition.

The target recovery hypothesis predicts an opposite result. Makovski and Jiang (2009) found that tracking performance increased when targets had feature identities but not when they had conjunction identities (compared to when all targets were identical), suggesting that conjunction identities cannot be used for recovering lost targets. At any rate, searching for a lost feature target should be substantially easier than searching for a lost conjunction target, especially a color \times color conjunction (Wolfe et al., 1990). Thus, if the repetition benefit is based on searching for lost targets, this benefit should be eliminated, or at least reduced, in the conjunction condition, compared to the feature condition.

Therefore, we have two contrasting predictions. According to the binding hypothesis, the repetition benefit should be larger in the conjunction condition. The target recovery hypothesis predicts the opposite: the repetition benefit should be observed in the feature condition, but not (or only minimally) in the conjunction condition.

Methods

Participants

Sixteen participants took part as paid volunteers. They ranged in the age from 18 to 50 years old (average 30.1 years), and all had normal or corrected-to-normal vision.

Stimuli

In both conditions, the stimuli had the same dimensions as in Experiment 1 (i.e., $2.37^\circ \times 2.37^\circ$). In the feature condition, there were eight different basic colors: red (0.641, 0.341, luminance: 11.5 cd/m^2), blue (0.148, 0.072, luminance: 4.82 cd/m^2), yellow (0.405, 0.521, luminance: 54.4 cd/m^2), green (0.293, 0.606, luminance: 26.7 cd/m^2), black (0, 0, luminance: 0 cd/m^2), white (0.289, 0.317, luminance: 60.6 cd/m^2), orange (0.482, 0.444, luminance: 24.0 cd/m^2), purple (0.277, 0.142, luminance: 4.36 cd/m^2). The background (in both conditions) was gray (0.282, 0.308, 10.8 cd/m^2). In the conjunction condition, there were 24 possible color \times color conjunctions. All conjunctions were two half-rectangles ($1.19^\circ \times 2.37^\circ$) put together to form a $2.37^\circ \times 2.37^\circ$ square. Each half-rectangle was made up of one of the basic colors of the feature condition. In order to ensure that no item would have a unique color with a given display, we restricted ourselves to a subset of the potential color combinations. The allowed conjunctions were red–blue, red–green, red–yellow, yellow–blue, yellow–green, green–blue, black–white, black–purple, black–orange, white–orange, white–purple, and orange–purple. Each conjunction could appear in two mirror-reversed versions (e.g., red–blue and blue–red were different stimuli). See Figure 5 for an example of the feature and conjunction stimuli we used.

Procedure

The experiment consisted of two sessions of 8 blocks of 30 trials each, for a total of 480 trials per participant. Each session took approximately 75 min, and the sessions could be performed on separate days. Within a session, all items were either defined by unique colors (*feature* session) or by color conjunctions (*conjunction* session). Each session comprised four unrepeated blocks and four repeated blocks, as defined in Experiment 1. Half of the participants started with the feature session, while the other half started with the conjunction session. Within each session, half of the participants started with the repeated condition, and the other half with the unrepeated condition. If a participant started with the repeated condition on the first session, she would also start with the repeated condition on the second session.

Results and discussion

Figure 6 plots identity and location accuracies as a function of repetition and session. A two-way ANOVA on identity accuracy with the factors repetition (unrepeated versus repeated) and stimulus type (feature or conjunction) as independent variables and identity accuracy as dependent variable revealed that there was a main effect of trial type ($F(1, 15) = 35.61$, $MSE = 71.96$, $p < 0.001$, $\eta^2 = 0.704$) as accuracy was higher in the feature condition, a



Figure 5. Example of the stimuli we used in Experiment 3.

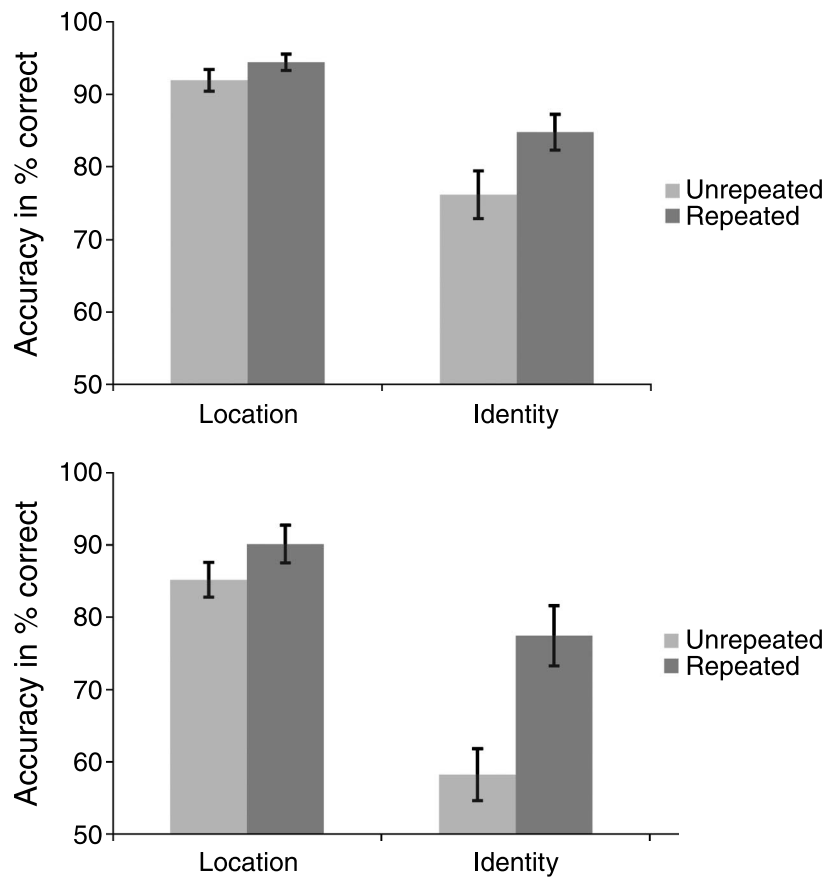


Figure 6. Location and identity accuracies in Experiment 3 in the unrepeated (light gray bars) and the repeated (dark gray bars) conditions. The top panel plots data from the feature task and the bottom panel plots data from the conjunction task.

main effect of repetition ($F(1, 15) = 58.06$, $MSE = 53.40$, $p < 0.001$, $\eta^2 = 0.795$) as accuracy was higher in the repeated blocks, and a significant interaction ($F(1, 15) = 5.27$, $MSE = 84.86$, $p < 0.05$, $\eta^2 = 0.260$) showing that the repetition benefit was larger for conjunctions than for features. The same pattern of results held for location accuracy, except now there was only a trend to a larger repetition benefit for conjunctions ($F(1, 15) = 2.91$, $MSE = 8.21$, $p = 0.11$, $\eta^2 = 0.163$). Furthermore, for both conjunctions and features, there was a significant repetition benefit, for both identity and location accuracies (Conjunctions: identity: $F(1, 15) = 36.50$, $MSE = 80.86$, $p < 0.001$, $\eta^2 = 0.709$; location: $F(1, 15) = 12.59$, $MSE = 15.54$, $p < 0.005$, $\eta^2 = 0.456$; Features: identity: $F(1, 15) = 10.39$, $MSE = 57.40$, $p < 0.01$, $\eta^2 = 0.409$; location: $F(1, 15) = 5.58$, $MSE = 8.96$, $p < 0.05$, $\eta^2 = 0.271$).

We found that location accuracy was significantly greater than identity accuracy in both conditions (Conjunctions: $F(1, 15) = 47.61$, $MSE = 131.86$, $p < 0.001$, $\eta^2 = 0.760$; Features: $F(1, 15) = 44.28$, $MSE = 58.29$, $p < 0.001$, $\eta^2 = 0.747$), demonstrating a binding problem even for simple color stimuli. We also replicated Makovski and Jiang's (2009) finding that tracking is better with feature stimuli than with conjunction stimuli.

Importantly, repetition improved tracking for conjunctions more than it improved tracking for features. This result directly contradicts the target recovery hypothesis and is in accordance with the binding hypothesis. Since the binding load in the feature condition is lower than in the conjunction condition, repetition has a larger effect in the latter case.

However, we could account for these results without the binding hypothesis. Instead of assuming that repetition directly improves target recovery, perhaps repetition improves memory for the target set. This in turn improves tracking performance because the better an observer knows the target set, the less likely she is to accidentally start tracking a distractor instead of a target. Thus, prevention of target–distractor swaps is more robust.

This hypothesis can explain why performance improves more for conjunction stimuli than for feature stimuli (since the latter are already very easy to remember, and there is little room for improvement). We designed Experiment 4 to test this hypothesis.

Experiment 4: A repetition effect without distractors

If the repetition benefit is due to improved prevention of target–distractor swaps, we should be able to eliminate it by eliminating the distractors. One advantage of using unique stimuli for tracking is that we can do that without making the task completely trivial (see Experiment 7 of Horowitz et al., 2007).

Methods

Participants

Eight participants took part as paid volunteers. Their ages ranged from 21 to 50 years (average 32.25 years), and all had normal or corrected-to-normal vision.

Stimuli and procedure

The stimuli and procedure were similar to those of Experiment 1, except that there were five targets and no distractors. The experiment consisted of 4 blocks of 30 trials and was performed without breaks. The entire experiment took approximately 45 min.

Results and discussion

The results are plotted in Figure 7. There was a significant repetition benefit ($F(1, 7) = 9.754$, $MSE = 5.41$, $p < 0.05$, $\eta^2 = 0.582$). Thus, the target–distractor swap hypothesis cannot completely explain the repetition effect. However, note that the repetition benefit seems to be reduced in comparison to the other experiments (in this experiment the repetition benefit is approximately 5%, whereas in the other experiments it hovers around 15%). This suggests that repetition does serve to reduce the probability of target–distractor swaps. Nevertheless, there remains a significant repetition benefit even when there are no distractors. How can we explain this?

Perhaps repetition improves tracking performance by reducing the memory load. There is certainly evidence that MOT and visual working memory are related, from both neural and behavioral sources (Allen et al., 2006; Drew & Vogel, 2008; Fougner & Marois, 2006; Howe, Horowitz, Morocz, Wolfe, & Livingstone, 2009; Oksama & Hyona, 2004). Repetition might make it easier to encode stimuli or to transfer stimuli from working memory to long-term

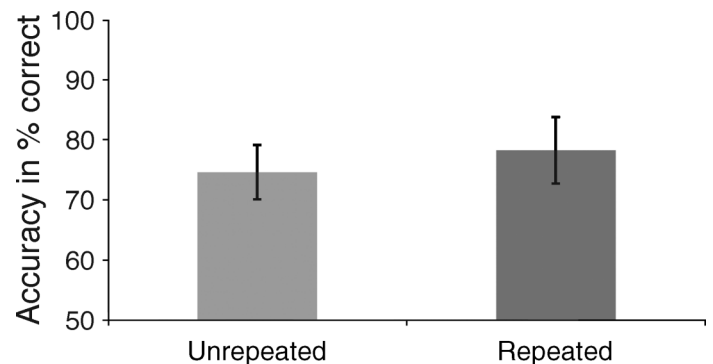


Figure 7. Accuracy in Experiment 4 in the unrepeated (light gray bar) and the repeated (dark gray bar) conditions.

memory, thereby freeing up resources that could then be devoted to the tracking task. We designed [Experiment 5](#) to test this hypothesis.

Experiment 5: Repetition improves identity tracking

In [Experiment 5](#), we contrasted the method used in the previous experiments (Horowitz et al., 2007; Makovski & Jiang, 2009; Oksama & Hyona, 2004, 2008), in which identities are constantly visible, with the original method of Pylyshyn (2004), in which identities are presented at the start of the trial and then masked during the motion phase. If repetition improves MIT via processes unrelated to the tracking of moving identities (such as reducing the memory load), then the repetition benefit should also show up in the masked condition. In contrast, if the repetition benefit derives primarily from processes involved in identity tracking, such as more efficient binding of identities and positions during tracking, then it should only be observed when the identities are visible during tracking.

Note that it is likely that identity accuracy will be greater in the visible identities condition, regardless of which account is correct. However, the memory load hypothesis predicts a larger repetition benefit in the masked condition, whereas the improved identity tracking hypothesis expects a larger repetition benefit when all items are visible during the motion phase.

Methods

Participants

Sixteen participants took part as paid volunteers. Their ages ranged from 21 to 50 (average 27.8) years, and all had normal or corrected-to-normal vision.

Stimuli and procedure

There were two conditions: *visible* and *masked*. The visible condition was identical to [Experiment 1](#). The

masked condition was similar except that during the motion phase, the items were replaced with the same masks used during the response phase (see [Figure 8](#)).

The experiment consisted of two sessions. Each session started with a practice block of 10 trials, followed by 4 blocks of 30 trials. Half of the participants started with the visible session in which the identities of all items remained visible throughout the trial, the other half of the participants started with the masked session, in which the identities were visible for 8 s before the start of the trial, but once the items started moving identities disappeared. Within each session, half of the participants started with the repeated condition, and the other half with the unrepeated condition. If a participant started with the repeated condition on the first session, she would also start with the repeated condition on the second session (and similarly if she would start with the unrepeated condition). The sessions were performed on different days. One session would be performed without breaks and took approximately 50 min.

Results and discussion

The data are presented in [Figure 9](#). We analyzed the results for identity and location accuracies separately. In both cases, a two-way ANOVA with factors condition (masked vs. visible) and repetition (repeated vs. unrepeated) revealed a main effect for condition (identity: $F(1, 15) = 54.43$, $MSE = 128.58$, $p < 0.0001$, $\eta^2 = 0.784$; location: $F(1, 15) = 51.96$, $MSE = 54.94$, $p < 0.0001$, $\eta^2 = 0.776$) as performance was higher when all items were visible, a main effect for repetition (identity: $F(1, 15) = 29.68$, $MSE = 42.00$, $p < 0.001$, $\eta^2 = 0.664$; location: $F(1, 15) = 18.72$, $MSE = 19.62$, $p = 0.001$, $\eta^2 = 0.555$) as performance was higher when the targets were repeated, and a significant interaction (identity: $F(1, 15) = 10.71$, $MSE = 50.74$, $p = 0.005$, $\eta^2 = 0.417$; location: $F(1, 15) = 6.70$, $MSE = 13.76$, $p < 0.05$, $\eta^2 = 0.309$) as repetition improved performance more in the visible than in the masked condition. Follow-up tests revealed that, for both identity and location accuracies, there was a repetition benefit in the visible condition ($F_s > 26.3$, $p_s < 0.001$, $\eta^2 > 0.64$), but no significant repetition benefit in the masked condition ($F_s < 2.55$,

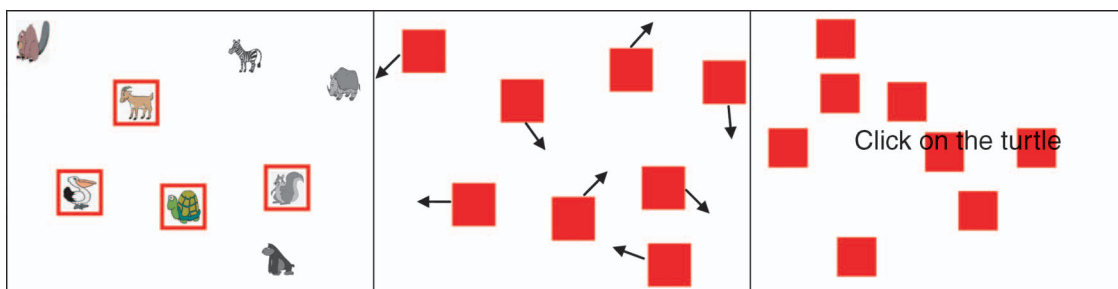


Figure 8. An example of the masked condition in [Experiment 5](#). First, the participant was presented with the targets, with their identities visible. At the start of the trial, all identities were masked. At the end of the trial, the participant indicated where a specific animal was.

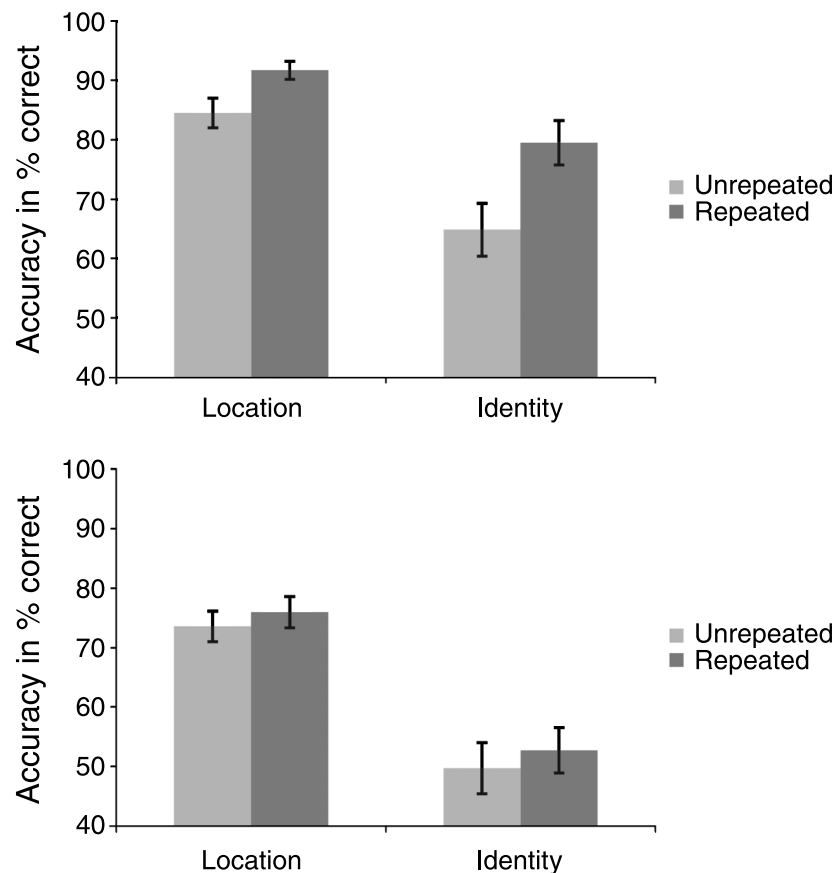


Figure 9. Location and identity accuracies for Experiment 5 in the unrepeated (light gray bars) and the repeated (dark gray bars) conditions. The top panel depicts accuracy in the visible condition and the bottom panel depicts accuracy in the masked condition.

$p_s > 0.13$, $\eta^2 < 0.15$). We then performed a two-way ANOVA on the results of the visible condition, with the factors condition (repeated or unrepeated) and accuracy (location or identity). This interaction contrast was significant ($F(1, 15) = 14.41$, $MSE = 15.43$, $p < 0.005$, $\eta^2 = 0.490$), indicating that the increase in identity accuracy was larger than the increase in location accuracy, thereby replicating the results of Experiment 1. A similar analysis of the masked condition revealed no significant effects. These data disconfirm the memory load hypothesis and are in line with the predictions of the binding hypothesis.

One problem with interpreting these results is that overall accuracy was lower in the masked condition. Since we do not know the shape of the accuracy \times difficulty curve for this task, it is possible, if not necessarily likely, that the masked condition lies in a domain where accuracy is less sensitive to difficulty than in domain of the visible condition data. Therefore, we ran a control experiment with eight paid volunteers, where we adjusted the speed in order to produce equivalent accuracy in both conditions. Observers ran in a block of 60 trials of the repeated condition in which speed was adjusted using a QUEST routine (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Watson & Pelli, 1983) to obtain 85% accuracy. We then ran all four conditions of the main experiment at this speed

(visible condition: mean = 13.7°/s (standard error = 2.4°/s), masked condition: 5.9°/s (1.1°/s)).

The control experiment replicated the results of the main experiment (see Figure 10). With regard to identity accuracy, we again found a significant interaction, indicating that the repetition benefit was larger in the visible than in the masked condition ($F(1, 7) = 6.99$, $MSE = 173.66$, $p < 0.05$, $\eta^2 = 0.500$), a repetition benefit in the visible condition, but not in the masked case ($F(1, 7) = 6.51$, $MSE = 426.59$, $p < 0.05$, $\eta^2 = 0.482$; and $F < 0.3$, $p > 0.6$, $\eta^2 < 0.05$, respectively). With regard to location accuracy, the data trended in the same direction, but none of the effects reached significance (interaction, indicating a larger repetition benefit in the visible condition: $F(1, 7) = 4.20$, $MSE = 3.49$, $p = 0.08$, $\eta^2 = 0.375$; visible condition, repetition benefit: $F(1, 7) = 2.136$, $MSE = 2.14$, $p = 0.19$, $\eta^2 = 0.234$; masked condition, repetition deficit: $F(1, 7) = 1.379$, $MSE = 0.87$, $p = 0.28$, $\eta^2 = 0.165$).

General discussion

In five experiments, we consistently found that repeating target identities improved both identity and position

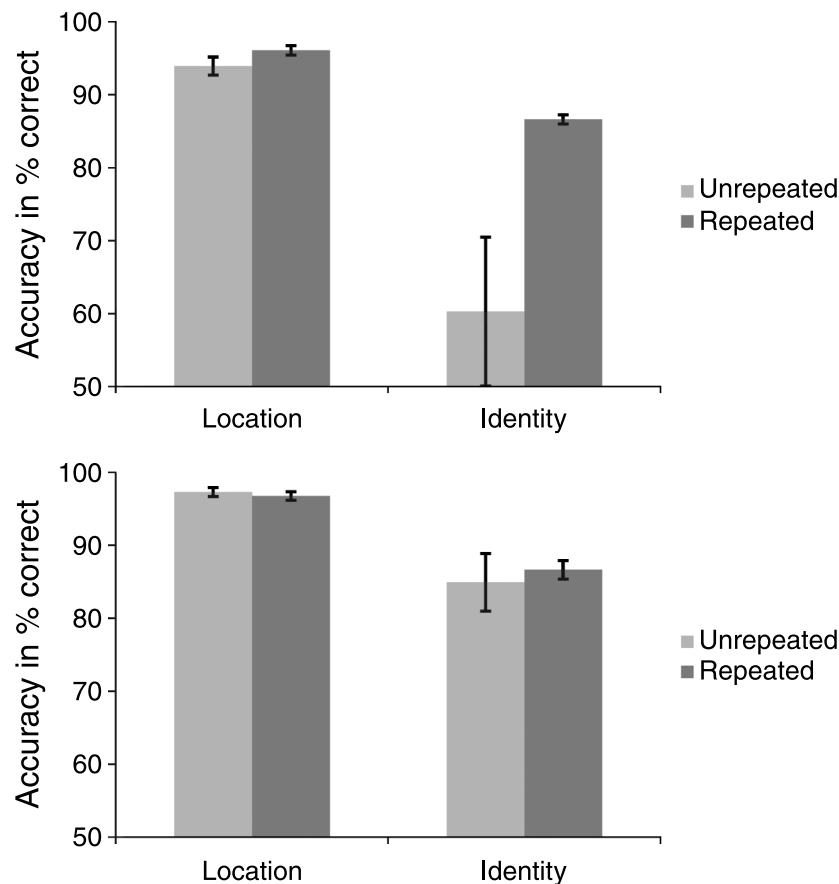


Figure 10. Location and identity accuracies for the control of [Experiment 5](#) in the unrepeated (light gray bars) and the repeated (dark gray bars) conditions. The top panel depicts accuracy in the visible condition and the bottom panel depicts accuracy in the masked condition.

tracking. In [Experiment 1](#), using cartoon animals, we found a solid repetition benefit. We found that identity accuracy was more than 15% higher with repeated than with unrepeated targets. [Experiment 2](#) showed that even when distractor and target identities were never swapped, performance was still better when targets were repeated. In [Experiment 3](#), we found that the repetition benefit was larger when the identities were defined by color–color conjunctions than if the identities were defined by unique colors. [Experiment 4](#) showed that the repetition benefit was still observed even when all stimuli were targets. Finally, [Experiment 5](#) showed that the repetition benefit is strongly reduced when target identities are not visible during the motion phase.

Four hypotheses

We proposed four hypotheses to explain the repetition benefit. First, in the repeated condition target and distractor roles are consistent across trials, whereas in the unrepeated condition, targets on one trial could become distractors on a subsequent trial and vice versa. This consistent mapping might account for the repetition benefit. [Experiment 2](#) compared the repeated condition to unrepeated conditions

with consistent mapping. In these unrepeated conditions, targets and distractors were selected from separate stimulus pools, assuring that targets could never become distractors (and vice versa). [Experiment 2](#) still showed better tracking performance in the repeated condition. Thus, consistent mapping cannot explain the entire repetition benefit.

Second, repetition might improve tracking by improving the ability to recover lost targets. This explanation is difficult to reconcile with the results of [Experiment 3](#), in which easy to search for feature stimuli show a smaller repetition benefit than difficult to search for color–color conjunctions. A modification of this hypothesis suggests that repetition allows observers to better prevent the swapping of a target for a distractor, because memory for the target set is improved. [Experiment 4](#) showed that this hypothesis cannot fully explain the repetition benefit, since a benefit was observed even in a target-only tracking task, in which the probability of swapping targets and distractors is nil. Since the benefit was reduced relative to the previous experiments, we can conclude that repetition does help reduce target–distractor swaps, but this cannot be the whole story.

Third, repetition might reduce the memory load, freeing up shared resources that could then be diverted to the tracking task. This hypothesis was tested in [Experiment 5](#)

by masking the objects during the moving phase. If repetition improves tracking by improving memory, then this manipulation should increase the repetition benefit, or at least leave it unchanged. However, the manipulation reduced the repetition benefit, disconfirming this account.

The fourth hypothesis, consistent with all of the results, is that repetition improves performance, because it improves processes of identity tracking, either through improving recognition and maintenance of target identities, or through facilitating identity–position binding during tracking. For example, if we assume that MIT proceeds via a cyclical updating of identity + position bindings (Oksama & Hyönä, 2008), the repetition of target identities might facilitate recognition of identities by changing the long-term representation of the object (Ellis, Shepherd, & Davies, 1979; Leveroni et al., 2000; Wolfe & Horowitz, 2004), reducing the time or neural resources needed to complete the identity–location binding operation on each cycle.

Summarizing, our data decisively show that repeating target identities increases performance on an MIT task. [Experiment 5](#) shows that this repetition benefit is mainly due to processes involved when the different identities are in motion. This strongly suggests that repeating target identities improves identity tracking. [Experiments 2, 3, and 4](#) are all consistent with the notion that this improved identity tracking is due to improved identity–location binding. However, our data do not exclude that other processes could also play a role.

How does identity repetition relate to “familiarity”?

In the current study, we have found that you are better at tracking a stimulus when you have had more exposure to this stimulus. In general, when you see a stimulus more often, you become more familiar with this stimulus, but the two terms are not interchangeable. For example, in [Experiment 1](#) of our study, participants may have become more familiar with the repeated cartoon animals than the cartoon stimuli in the unrepeated condition. However, it would be hard to defend the proposition that our participants were truly “unfamiliar” with a blue square simply because it was presented in the unrepeated condition of [Experiment 3](#). On the other hand, while not all objects you have seen more often are more familiar, the reverse does hold. You have to have seen an object (or something similar to it) repeatedly before it could be considered familiar to you. This assertion is congruent with the findings of Oksama and Hyönä (2008). They compared tracking familiar to tracking unfamiliar objects (where familiar objects were real objects and unfamiliar objects were pseudo-objects) and observed an advantage for tracking familiar objects. They also replicated the finding using face stimuli where the “familiar” faces were famous faces (e.g., Albert Einstein and Saddam Hussein) likely to be well known

to their observers. Recall that in the [Introduction](#) section we outlined that several mechanisms may have caused the observations of Oksama and Hyönä (2008), such as their familiar objects being more nameable. We speculate that the findings in both studies are caused by the same underlying mechanism, namely improved identity tracking (i.e., improved maintenance and recognition of identities or improved identity–location binding) of familiar (or repeated) objects.

Two systems or limited resources

One way to summarize our results is that tracking identities in addition to locations is effortful (because repeating identities makes it easier) and pulls some resources away from tracking locations alone (because we get significant increases in both location and identity accuracies).

These results rule out a simple class of models of multiple identity tracking in which identities come along for free as long as locations are tracked. This model, for example, was assumed by Pylyshyn (2004) when he designed his studies. It also underlies the “impletion” idea behind *object file* theory (Kahneman et al., 1992). The object file theory suggests that directing attention to an object creates a file that contains all relevant data regarding this object, including its position. So the object file theory is equivalent to models where identities come along for free, but now, after identities have been acquired, locations get a free ride. In neither case would it be possible to trade off performance between identity and location tracking.

The fact that identity tracking is effortful (for instance because it requires updating position and identity bindings) is consistent with either a unitary account of MIT, in which the same neural system is responsible for tracking positions and identities or a dual account such as that proposed by Horowitz et al. (2007). However, if position tracking and binding are carried out by separate systems, then making binding easier should not affect position tracking. In all of our experiments, improved identity accuracy was accompanied by improved location accuracy. Thus, our data support a unitary account of MIT.

Our data are furthermore consistent with a limited resource model. If we assume that tracking positions and binding identities to positions are processes that share some neural resource, then if repeating targets makes binding more efficient, some resources are freed up for tracking, thus improving both identity and location performances. A limited resource model of MOT has been proposed by Alvarez and Franconeri (2007). According to this model, tracking is accomplished via an array of flexibly allocated indexes (or FLEXes), such that when few targets are tracked, they receive many indexes, and when many targets are tracked, each target receives fewer indexes (the FLEX model can also be implemented in a serial fashion). Thus, tracking performance degrades smoothly as tracking

load increases. However, under this account, the limited resource is really spatial resolution: when few targets are tracked, each target receives more resources (FLEXes), and the resolution increases so that it can be tracked at a higher speed. In order to explain our data, we have to add the assumption that the visual system can trade off spatial resolution for identity information (speculatively, in the form or featural resolution), so that when recognizing objects becomes easier, spatial resolution improves. Thus, these data suggest a more general limited resource account of MOT and MIT.

One way to understand this trade-off might be to assume that we can trade off spatial resolution for resolution in a non-spatial dimension. For example, if observers track colored squares, making the tracking task more difficult might reduce the resolution of color information (see Zhang & Luck, 2008 for a method to measure color resolution).

Another approach is to assume that MIT consists of a series of update cycles. On each cycle, the system updates both the positions and the bindings. If binding is more efficient, the update cycles can proceed at an increased rate, leading to reduced position error, such as in the MOMIT model (Oksama & Hyönä, 2008).

Conclusions

Performance in a multiple identity tracking task improves when targets are repeated often, provided the targets are visible when they are in motion. This effect is most likely due to identity tracking improving because of target repetition. Our results indicate that tracking positions and identities are not independent tasks. Our data thus argue against simple models in which identity comes for free with position tracking and also against dual-system models in which position tracking and identity tracking are accomplished by different neural hardware. Instead, we suggest that a limited resource model, in which position tracking and identity binding compete for the same resource pool, may explain tracking.

Appendix A

Figure A1.



Figure A1. The cartoon animals used in Experiments 1, 4, and 5.

Acknowledgments

We would like to thank Shin'ya Nishida and two anonymous reviewers for their constructive comments on an earlier version. This research was funded by a Rubicon grant from NWO (Netherlands Organization for Scientific Research) to Ya'ir Pinto, and NIH grant MH 65576 to Todd S. Horowitz.

Commercial relationships: none.
Corresponding author: Ya'ir Pinto.
Email: yair.pinto@gmail.com.

Address: 64 Sidney Street, Suite 170, 02139 Cambridge, MA, USA.

References

- Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology, 18*, 337–347.
- Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2006). Multiple-target tracking: A role for working memory? *The Quarterly Journal of Experimental Psychology, 59*, 1101–1116.
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you attentively track?: Evidence for a resource-limited tracking mechanism. *Journal of Vision, 7*(13):14, 1–10, <http://www.journalofvision.org/content/7/13/14>, doi:10.1167/7.13.14. [PubMed] [Article]
- 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. *Behavioral and Brain Sciences, 24*, 87–114.
- Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *Journal of Neuroscience, 28*, 4183–4191.
- Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*, 431–439.
- Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory. *Psychological Science, 17*, 526–534.
- Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic Bulletin & Review, 15*, 802–808.
- 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.
- Howe, P. D., Horowitz, T. S., Morocz, I. A., Wolfe, J. M., & Livingstone, M. S. (2009). Using fMRI to isolate components of the multiple object tracking task. *Journal of Vision, 9*(4):10, 1–11, <http://www.journalofvision.org/content/9/4/10>, doi:10.1167/9.4.10. [PubMed] [Article]
- 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.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*, 171–216.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.
- King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research, 34*, 885–912.
- Kroll, J., & Potter, M. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning & Verbal Behavior, 23*, 39–66.
- Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience, 20*, 878.
- Makovski, T., & Jiang, Y. V. (2009). Feature binding in attentive tracking of distinct objects. *Visual Cognition, 17*, 180–194.
- Oksama, L., & Hyona, 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.
- Pinto, Y., Olivers, C. N., & Theeuwes, J. (2005). Target uncertainty does not lead to more distraction by

- singletons: Intertrial priming does. *Perception & Psychophysics*, 7, 1354–1361.
- 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: I. Tracking without keeping track of object identities. *Visual Cognition*, 11, 801–822.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179.
- Saiki, J. (2002). Multiple-object permanence tracking: Limitation in maintenance and transformation of perceptual objects. *Progress in Brain Research*, 140, 133–148.
- Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, 3(1):2, 6–21, <http://www.journalofvision.org/content/3/1/2>, doi:10.1167/3.1.2. [PubMed] [Article]
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66.
- 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.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 37, 571–590.
- Tipper, S. P., Weaver, B., & Wright, R. D. (1998). *The medium of attention: Location-based, object-centered, or scene-based?* New York, NY: Oxford University Press.
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14, 411–443.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113–120.
- Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron*, 24, 11–17, 111–125.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495–501.
- Wolfe, J. M., Yu, K. P., Stewart, M. I., Shorter, A. D., Friedman-Hill, S. R., & Cave, K. R. (1990). Limitations on the parallel guidance of visual search: Color \times color and orientation \times orientation conjunctions. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 879–892.
- 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.