Vision Research 145 (2018) 49–55

Contents lists available at ScienceDirect

Vision Research

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

# Event monitoring: Can we detect more than one event at a time?

Chia-Chien Wu[a,b,*], Abla Alaoui-Soce[b], Jeremy M. Wolfe[a,b]

[a] *Harvard Medical School, United States*
[b] *Brigham & Women's Hospital, Boston, MA, United States*

ARTICLE INFO

Number of reviewers = 2

*Keywords:*
Multiple object tracking
Attention
Change detection

ABSTRACT

A prior study by Wu and Wolfe found that the capacity for event monitoring (e.g. did an item change its state?) is more limited than for classic multiple object tracking. That limited capacity, K, could arise from either of two situations. It could be that people can detect K events simultaneously or it could be that they can successfully detect just one event at a time while monitoring K out of a total of N items. In the three different experiments of the present study, observers were asked to monitor a set of moving objects while watching for two critical events occurring in that set. Observers' performance can be well described by a model that includes an ability to detect two changes at once. Our results suggest that the capacity for event monitoring is further limited when tracking an additional event, but within the monitored set, people can detect at least two events simultaneously.

## 1. Introduction

In a surveillance task, the success of detecting the potential threat depends not only on how many people in the crowd a security guard can watch at the same time, but also on how well suspicious behaviors among the monitored agents can be detected. Wu and Wolfe (2016) conducted a series of experiments and asked their observers to track a group of entities and watch for a specific event in that group. They found that people could only track a very limited set of items when their task was to detect an event during a sustained monitoring task. This "event monitoring capacity", *K*, is significantly smaller than the position tracking capacity, measured in conventional Multiple Object Tracking (MOT) tasks (Pylyshyn & Storm, 1988). Typical MOT capacity is 3–4 items though it varies with the specific task (Alvarez & Franconeri, 2007). Event monitoring capacity is around 2–3 items. The nature of this event monitoring capacity is not entirely clear. In MOT, the tracking capacity is usually thought to represent the number of objects that can be tracked concurrently. Similarly, the capacity of the Multiple Identity Tracking (MIT) task represents the number of objects' whose identity can be addressed during the position tracking (Horowitz et al., 2007; Makovski & Jiang, 2009; Oksama & Hyönä, 2004). Does the Multiple Event Monitoring (MEM) capacity measure the number of events people can detect simultaneously during tracking? Can observers detect two events at the same time? Alternatively, the MEM capacity could represent the size of the subset of items that can be monitored for an event during tracking. That is, observers might be able to keep track of the locations of, say, 2–3 items as shown in MEM limit, but they might be further limited to noticing a single change to those items.

Even detecting a single event requires observers to encode the initial states of tracked agents, so that they can detect any state change once it has happened. The change blindness literature suggests that when viewing a scene, the visual information that is actually available to support change detection is much more limited than what we naively believe that we see (Simons & Levin, 1997; Simons & Rensink, 2005). Rensink (2000) proposed his "coherence theory" to explain how changes can be perceived even if only a little information is encoded. In his model, during the early visual process, a low-level prototype object (or proto-object) is formed across the visual field and a small subset of these prototype objects would be attended to create a single higher-level structure, a *nexus*. The objects in this nexus form a coherence field over space and time. A change can be detected only if it occurs to an object held by focused attention in that nexus. Moreover, since the information about the attended objects is pooled into the single nexus, it is not possible to distinguish whether a detected change is the result of a single change signal or multiple change signals. If the attention in a sustained monitoring task operates in the way that coherence theory describes, the event monitoring capacity *K* might represent the number of proto-objects people can attend to simultaneously. However, while observers might be able to detect any change in that group of objects, coherence theory would seem to suggest that they would not be able to differentiate between one or several changes in that group.

In an alternative to the coherence theory account, multiple event monitoring might operate in a manner similar to multiple object tracking where each individual object in a limited set can be tracked in parallel. Howe, Cohen, Pinto, and Horowitz (2010) tested observers in two tracking conditions. In one condition, all items moved then all

* Corresponding author at: Brigham & Women's Hospital, Boston, MA, United States.

https://doi.org/10.1016/j.visres.2017.10.009
Received 18 July 2017; Received in revised form 14 September 2017; Accepted 17 October 2017
Available online 31 March 2018

stopped simultaneously. In the other condition, only half of the items moved. When they paused, the other objects moved so that, at any given time, only half the objects were moving. If the tracking was completed in series, observers should perform better in the sequential condition, where only half the objects would need to be tracked at any one moment, than in the simultaneous condition where all the targets have to be tracked during each moving phase. However, Howe et al. (2010) found that the tracking performances were similar between the sequential condition and the simultaneous condition, which suggests that multiple object tracking was operating in parallel over the whole set. Other studies also found a similar parallel operation across multiple moving objects (Alvarez & Cavanagh, 2005; Störmer, Winther, Li, & Andersen, 2013). Thus, if event monitoring and position tracking operate similarly, it should be possible for more than one event to be detected in parallel. The MEM capacity, *K*, may represent the number of events people can detect at the same time during the tracking. To test this possibility, we conducted MEM tasks in which two events either occurred at the same time or occurred asynchronously. To preview our results, we found that though a second event affects performance relative to detection of a single event, people could monitor for two simultaneous events just as well as for sequential events.

## 2. Experiment 1

In Experiment 1, we conducted a similar event monitoring experiment to Wu and Wolfe (2016) using photorealistic objects selected from Brady, Konkle, Alvarez, and Oliva (2008). Each object had two different states and could change from one state to the other (e.g. an open book becomes a closed book). Critically, in the new experiment, there were two target events instead of one. Two objects could either change their states at the same time, or at different times. If observers are able to detect two events at the same time, an interval between state changes should not affect the monitoring performance. On the other hand, if observers could only notice one change at a time, then performance would be worse when the two changes happen simultaneously than when they occur sequentially.

### 2.1. Method

#### 2.1.1. Participants

Twelve participants (8 female, average age 24) recruited from the Brigham and Women's Hospital's volunteer pool took part in Experiment 1. All participants gave informed consent and were compensated $10/hour for their participation. The informed consent was approved by the Partners Human Research Committee. All participants passed the Ishihara test for color blindness and had normal or corrected-to-normal vision.

#### 2.1.2. Apparatus and stimuli

Stimuli were displayed on a 24″ screen (iMac model A1225) with a resolution of 1920 × 1200 pixels. All items moved within a 20° × 20° imaginary window at a viewing distance of 60 cm. The experiments were run using MATLAB 8.3 with Psychtoolbox (Brainard, 1997; Kleiner et al., 2007). On each trial, all items were randomly chosen from a set of 31 different objects and each of these objects has two distinct states (e.g. in Fig. 1, a book can be open or closed). All items were presented on a white background with a size of about 1.89° × 1.89°.

#### 2.1.3. Procedure

Experiment 1 consisted of 3 different set sizes (4,6,8) and 2 change time conditions (*same* or *different* change time). Thus, there was a total of six blocks with 50 trials each. The order of blocks was counterbalanced. On each trial, all N objects would first appear and remain stationary for N seconds so that observers had enough time to encode objects' identities and their initial states. All objects then began to move within an 20° × 20° imaginary window and the movement velocity was set to 4°/s. If two objects travelled across each other's paths, one would opaquely occlude the other. In the *same time* condition, both targets would simultaneously change their states at a time point randomly chosen from the interval between the 2nd and 6th second after motion started. In the *different time* condition, both targets would change their states at two different time points selected from the same range. The average time interval between two changes was about 1.6 s.

To prevent any attention-grabbing pop-out effect that might be caused by the target events, in addition to moving along with its own path, each item would also simultaneously rotate 30°in one direction for 250 ms and then return to its original orientation. This produced transients that were not associated with state-changes. Observers were informed about the identity of the time condition block (same/different) that they were running. They were told that the goal of the task was always to find the two target events (the two objects that changed states). They would press a key to stop the movement ending the trial after they found one or two targets. Once the observers ended the trial, the items would stop moving and be replaced by empty squares. The observers were asked to indicate the locations of both targets by mouse click. A trial would be counted as a miss and automatically terminated if no response was made within two seconds after the second state change occurred. Feedback was given after the response was made. Note that, though observers were asked to find both targets, they were not constrained to make a response only after the second target was detected. Therefore, in principle, in the different time condition, observers could make a response before the second event occurred and guess about the location of the second event.

#### 2.1.4. Results

There are two questions of interest here. First, can observers detect two events that occur at the same time and, second, how many items can be monitored concurrently (the MEM tracking capacity)? As shown in Fig. 2, tracking accuracy decreased when the set size increased (Two-way repeated measures ANOVA $F(2,22) = 228.06$, $p < 0.001$, $\eta_p^2 = 0.95$). If observers could only detect one change at a time, the performance in the same change time condition should be markedly worse than in the different change time condition, because they would always miss at least one of two targets even if both targets were concurrently tracked. The critical observation is that the performances were quite similar between the two conditions. In fact, it is the *different* condition that appears to be marginally worse in a standard ANOVA (Two-way repeated measures ANOVA $F(1,11) = 3.81$, $p = 0.08$, $\eta_p^2 = 0.26$). Though a Bayesian repeated measures ANOVA favors the null by 3×. Overall, it appears that observers can detect two changes at the same time in this task. There is no indication that performance is reduced in the *same* condition.

#### 2.1.5. MEM capacity analysis

To estimate the MEM capacity, we first need to determine the probabilities of having at least one target reside within the monitored subset. To analyze the possibilities, let us assume that observers are monitoring the states of K out of the total of N items in the display; thus, they can detect changes in the K-item subset, but will miss changes (or, possibly guess about targets) in the remaining N-K items.

With two targets among N total objects, there are three possible outcomes when observers monitor K items:

(1) Both targets are in the subset;
(2) Only one of two targets is in the subset;
(3) Neither target is in the subset.

Because of the 2-s response deadline, only the first two options can lead to correct detection of at least one event before the deadline. If both targets are in K, both targets could be correctly detected and

**Fig. 1.** The experimental procedure in Experiment 1. All N items would start to move after remaining stationary for N seconds. Two objects would change their states either at the same time or at different times. The times for state changes were randomly chosen between the 2nd and 6th second (as shown in the figure, the closed book became an open book, and the open toilet became a closed toilet). Observers were told to press a key as soon as they detected both events. After observers made the response, each item would be replaced by an empty square and observers would need to indicate the locations of two events by mouse click.

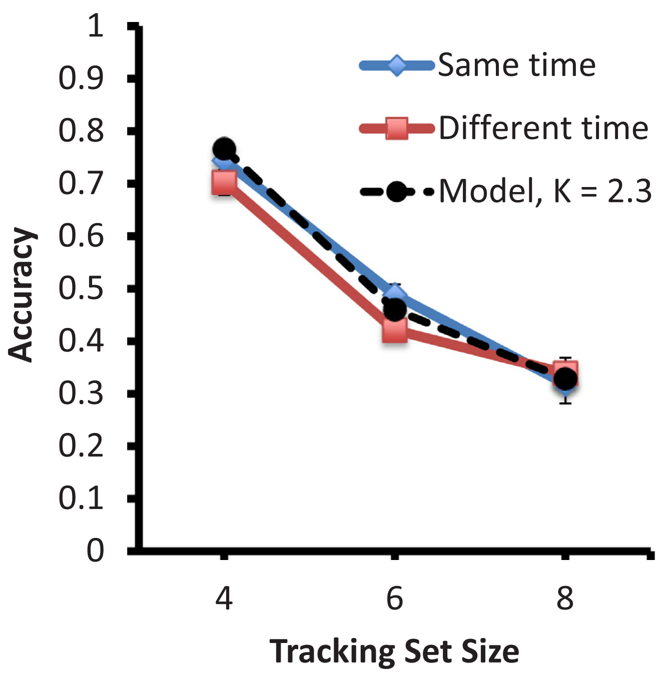


**Fig. 2.** Tracking accuracy in Experiment 1. The blue solid line shows the performance in the same change time condition. The red solid line shows the performance in the different change time condition. The dashed line indicates the estimation from the Model. Error bars are ± 1 SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

located before the deadline. If one target is in K, observers will detect that change and could choose to guess about the other. This would be especially true in blocks of simultaneous changes. If you detect one change and you know there were two changes, you might as well push the button and guess about the second item. If neither target is in K, observers would not be able to detect any target in time and a correct guess would require that the observer guess the time and the locations of the changes.

For a display with N items, there are $\binom{N}{K}$ ways to choose a K item subset. Out of these possibilities, we can calculate the proportion that would include both targets in that subset, K. The number of ways to pick both targets as two objects within the K item subset is $\binom{2}{2} = 1$. When both targets are selected, the number of remaining spaces in K becomes *(K-2)* and the number of remaining objects that could be placed in the subset is *(N-2)*. Thus, the number of combinations that include both targets in K among N objects is $\binom{2}{2}\binom{N-2}{K-2}$ and the probability of getting both targets is this number over the number of all possible combinations:

$$p1 = \frac{\binom{2}{2}\binom{N-2}{K-2}}{\binom{N}{K}} = \frac{K*(K-1)}{N*(N-1)} \tag{1}$$

To calculate the probability that one and only one target is in subset K, note that there are $\binom{2}{1} = 2$ ways to pick one target. The number of remaining spaces in K will be *(K-1)*. The number of remaining objects which can be selected into tracked items is still *(N-2)* because the other target cannot be selected into the subset. Therefore, the number of combinations having one and only one target is $\binom{2}{1}\binom{N-2}{K-1}$ and the probability of only one target selected can be calculated as

$$p2 = \frac{\binom{2}{1}*\binom{N-2}{K-1}}{\binom{N}{K}} = 2*\frac{K*(N-K)}{N*(N-1)} \tag{2}$$

As noted, when no target is selected into the K-item subset, the observer will not respond and the trial would be terminated two seconds after the last change occurs.

Given these probabilities, we can estimate the value of K that represents the number of items that the observer can successfully monitor for a state change. Since the *same time* and *different time* conditions produce similar results, for simplicity, we used the data from the *same time* condition to estimate K. The trials where observers responded before any change occurred were excluded (~10.1%).

The values plotted in Fig. 2 are the proportion of total targets that were accurately identified. This is composed of cases where both targets were in subset K, occurring with probability = *p1*. Since you get two targets for each of these cases, the expected contribution of these cases to the total would be $p1 * 2$. When only one target was included in the tracked set and the other target was in the untracked set (*p2*), observers will detect that one target from the K tracked items. They will guess the other target by picking on item from the remaining N-K untracked items. The probability that this is correct is $\left(\frac{1}{N-K}\right)$. Therefore, the expected number of targets contributed in this condition would be $p2*\left(1+\frac{1}{N-K}\right)$. No target will be collected when no targets are included in K. There will be no occasion to guess and observers are assumed to miss both targets. Thus, the total number of selected targets can be summarized as:

$$T_{selected} = p1*2 + p2*\left(1+\frac{1}{N-K}\right) = \frac{2K}{(N-1)} \tag{3}$$

In multiple event monitoring, the performance *P* is computed by the number of events correctly detected divided by the total number of target events in the trial. Therefore, when there are two target events in a trial, $P = \frac{T_{selected}}{2}$ and the performance of 0.5 would be assigned on a trial when only one of the two targets was correctly detected.

We can replace $T_{selected}$ with $2P$ in formula 3 and derive the capacity K:

$$K = P*(N-1) \tag{4}$$

By using Eq. (4), the capacity estimation was about 2.3 items, which is similar to the result in Wu and Wolfe (2016). These results suggest that the MEM capacity $K$ is not much greater than 2 but, if observers happen to be monitoring the correct items, two events can be detected, even if they occur at the same time.

## 3. Experiment 2

The results of Experiment 1 show that it is possible for observers to detect more than one event at a time during multiple event monitoring. Though K, the estimated capacity, is similar to what has been found in our previous event monitoring studies, it remains unclear whether adding an extra event imposes a cost on the event monitoring capacity. In order to investigate how an additional event could affect MEM performance, in Experiment 2, observers conducted MEM tasks that contained either a single event or two concurrent events.

Performance in Experiment 1 might also have been affected by the observers' ability to encode the initial states of objects. That is, the estimated capacity might measure, not only the ability to detect the state changes when they occurred, but also the ability to encode the objects' initial states and to notice that target objects were now in different states, even if observers did not detect the changes at the moment when the events occurred. Thus, in Experiment 2, we minimized the memory component by having observers track a set of identical items and watch for a "drop-off" event among them.

### 3.1. Method

#### 3.1.1. Participants

Twelve observers (6 female, average age 29) participated in Experiment 2. All observers were screened with the same procedures as Exp 1 and were compensated $10/hour for their participation.

#### 3.1.2. Apparatus and stimuli

The apparatus used in Experiment 2 was identical to Experiment 1 but the stimuli were somewhat different. In Experiment 2, the tracked items were all identical dark gray circles with a diameter of 1.2° as shown in Fig. 3. Each target circle (target set size N = 4, 6, 8) was attached a small black circle with a diameter of 0.5° as a simulation of an agent holding a bag. There were another four distractor circles that did not have attached black circles. In addition, there were four static black circles randomly placed in the display. We will refer to the black circles as 'bags' that can be dropped and we will call the larger moving circles 'agents'. In total, there were (N + 4) agents and four static bags in the display.

Every target agent held a bag at the same position in a given trial. As shown in Fig. 3, the bags were attached to the lower right side of each tracked agent. This position of the attached bag was varied each trial.

#### 3.1.3. Procedure

As the objects moved about the screen, observers were asked to detect one or two simultaneous events where the target agent(s) would "drop" its bag on the ground. Observers participated 6 blocks with three different tracking set sizes (*N*) 4, 6 or 8 and two tracking event types (one drop off or two simultaneous drop offs). Each block contained 50 trials and a total of 300 trials were tested. Note that observers only had to track the items with a bag (N) rather than tracking all items (N + 4) on the screen.

On each trial, all *N* objects would initially appear and remain stationary for *0.5* * *N* seconds. All objects then began to move in straight lines within a 20° × 20° imaginary window. At a time randomly selected between the 2nd and 6th seconds after the start of motion, one or two target discs (depending on the block) would drop their bags on the ground, then keep moving. Observers were required to press a key as soon as they detected the drop-off event(s). Since the different numbers of drop-off events were run in separate blocks, observers knew how many targets needed to be detected in the current condition. The order of tracking set size and the number of drop-off events were counterbalanced. Similarly, if a trial was not responded within 2 s after the event(s) occurred, the trial would be terminated and counted as a miss error. Once observers made their response, all items on the screen became identical circles (i.e. all bags would be removed) and observers would have to select which item(s) was the target circle with mouse click.

#### 3.1.4 Data analysis

Similar to Experiment 1, the tracking performance was calculated as the number of targets successfully selected divided by the total numbers of targets. As before, we assume the target event would be detected when the event occurred to an object in the monitored subset of items. Trials where the response was made before the event occurred were excluded (~4.7%). When there is only one target on a trial, no guessing strategy can be used because of the 2 s response deadline, which makes successful guessing unlikely. Therefore, the tracking accuracy, P, is also given by the number of items observers can monitor (K) divided by the tracking set size in the display (N). Thus, the capacity, K, can be estimated as

$$K = P*N \tag{5}$$

For those blocks where there were two target events per trial, the estimate of capacity, K, is the same as in the equivalent condition of Experiment 1 and is given by equation 4 ($K = P*(N - 1)$) as

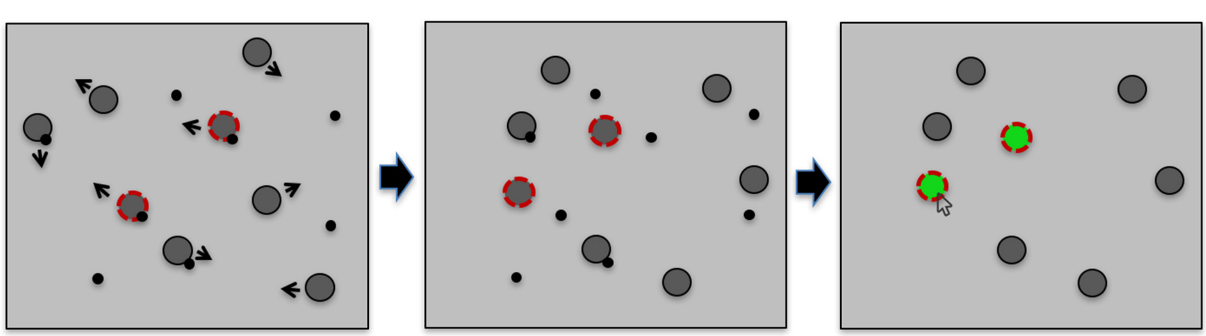

**Fig. 3.** The stimuli and procedure used in the 2-target condition of Experiment 2. All N items remained stationary for 0.5 * N seconds then started to move. Two targets would drop their bags simultaneously. Observers were asked to make their response as soon as they detected both events then use the mouse to indicate the target locations.

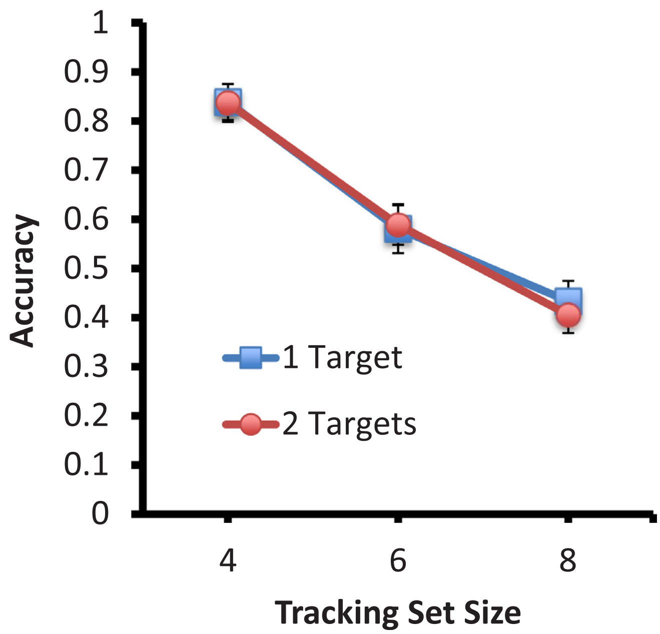


**Fig. 4.** Tracking Accuracy in Experiment 2. The blue solid line shows the performance in the 1-target condition and the red line shows the performance in the 2-target condition. Error bars are ± 1 SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

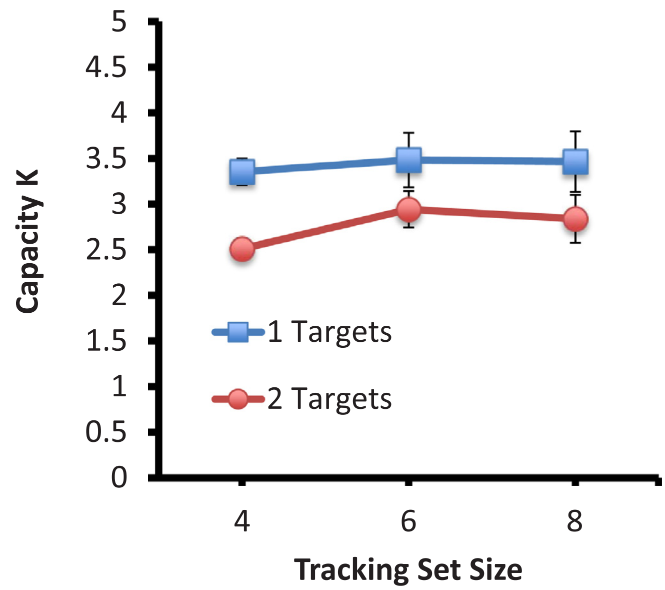


**Fig. 5.** Estimated capacity in Experiment 2. Error bars are ± 1 SEM.

mentioned above.

#### 3.1.5. Results

Fig. 4 shows the tracking accuracy as a function of tracking set size. As expected, the accuracy decreased with set size (Two-way repeated measures ANOVA $F(2,22) = 59.02$, $p < .001$, $\eta_p^2 = 0.84$). Interestingly, the tracking performance in the 2-target conditions was not worse than the performance in the 1-target condition, $F(1,11) = 0.003$, $p = 0.96$). The Bayes factor calculation also favors the null hypothesis by 4×.

Notice that, unlike in the one-target condition in which there was no guessing component, observers in the 2-target condition can always make a guess about the second target, so long as they correctly detected the first target event in time. This means that we would expect performance to be somewhat better in the two target case, if the monitoring capacity was the same in the two conditions. Alternatively, if the performance is, in fact, the same in the two conditions, we would need to conclude that monitoring capacity was reduced in the two target condition. This can be seen in Fig. 5 where the two-target capacity is computed from Eq. (4) with its guessing component and the one target capacity is computed from Eq. (5) without guessing.

When there was only one target event, the average tracking capacity was about 3.4. When there were two target events during the MEM, the average capacity decreased to about 2.7 ($F(1,11) = 21.7$, $p < 0.001$, $\eta_p^2 = 0.66$). There was no effect on tracking set size ($F(2,22) = 1.01$, $p = 0.38$, $\eta_p^2 = 0.08$). No interaction was found between the number of events and tracking set size ($F(2,22) = 0.35$, $p = 0.71$, $\eta_p^2 = 0.03$). This result shows that, if we assume guessing in the two-target case, observers' monitoring capacity is reduced by the need to monitor for two events. If we were to assume no contribution from guessing during the two event block, then Eq. (5) would govern both one- and two-event conditions and estimated capacity – like measured performance – would be essentially the same in the two conditions. In either case, the data indicate, as in Experiment 1, that observers are able to simultaneously detect two events.

## 4. Experiment 3

The results in Experiment 2 show that even if the estimate of monitoring capacity is somewhat reduced when there are two events, observers seem to be capable of detecting both of those events at the same time. In Experiment 1, it could have been that similar performance in same and different time conditions was achieved by a serial detection mechanism. Observers might have only detected one change then kept looking for the other item that was in a changed state since they knew that a second change had occurred. In this case, if an observer had not seen the second target in the currently tracked set, that observer might even search for that second, changed target amongst items that were not previously tracked. This is also possible in Experiment 2 when observers are asked to look for simultaneous changes. In Experiment 3, we attempt to thwart this strategy by mixing same and different time trials in the same block. This reduces the chance that a second target event could be found by a strategy of detecting one change and inferring the second by locating an item whose state had changed. Suppose that an observer noted one change. A simple search for another item that had changed state might not be successful since an item, that was unchanged when the observer examined it, could change state at a later time. Mixing simultaneous and sequential trials reduces the utility of looking around for an item that is in a new state. At the very least, mixing sequential and simultaneous conditions would force observers to repeatedly check for such an item.

### 4.1. Method

#### 4.1.1. Participants

Thirteen participants (1 male, average age 28) were tested in Exp 3. All were screened with the same procedure as before and were compensated $10/hour for their participation.

#### 4.1.2. Apparatus and stimuli

The apparatus and stimuli used in Experiment 3 were same as those used in the two-target condition of Experiment 2, except that the time interval between two drop-off events was either 0, 200 ms or 1000 ms.

#### 4.1.3. Procedure

The procedure used in Experiment 3 was identical to those used in the 2-target condition of Experiment 2 except that, on each trial, the interval between the two drop-off events was randomly selected from

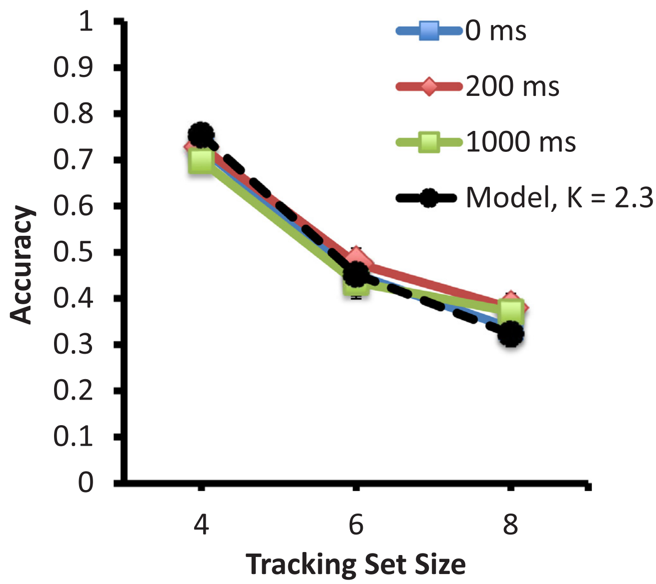


**Fig. 6.** Tracking accuracy in Experiment 3. The solid lines show obervers' performances for the time intervals of 0 (blue), 200 ms (red) and 1000 ms (green). The black dashed line shows the estimated performance from the Model. Error bars are ± 1 SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

one of three time intervals (0, 200 ms, 1000 ms). Each monitoring set size (N) 4, 6 or 8 was tested in two blocks of 50 trials and a total 6 blocks were tested. The order of set sizes was counterbalanced.

#### 4.1.4. Data analysis

Similar to previous experiments, we excluded the trials in which observers made a response before any event happened (~6.2%).

#### 4.1.5. Results

Fig. 6 shows performance for each of the three time intervals between two events in each monitoring set size. As expected, the monitoring accuracy decreased as a function of set size (Two-way repeated measures ANOVA: $F(2,24) = 110.24$, $p < .001$, $\eta_p^2 = 0.90$). Critically, the different time intervals between the two events did not affect the tracking performance (Two-way repeated measures ANOVA: $F(2,24) = 2.38$, $p = 0.11$, $\eta_p^2 = 0.17$). The Bayes factor calculation also favors the null hypothesis by 9×. No interaction was found between tracking set size and time interval ($F(4,48) = 1.77$, $p = 0.15$, $\eta_p^2 = 0.13$). The similar monitoring performances across different time intervals suggest that observers were able to monitor two events at the same time, such that the relative timing of the second event had no effect on the monitoring performance. The monitoring performance can be well described by the Model as explained in Experiment 1 and produces an estimated capacity of 2.3 items.

## 5. General discussion

Wu and Wolfe (2016) found that the capacity to monitor events in a dynamic display is quite limited. This MEM capacity appears to be significantly smaller than the position tracking capacity measured in a standard MOT task. It was not clear, from this previous work however, whether the MEM capacity, K, simply represented a limit on the number of items being tracked or a limit on the ability to detect multiple changes at the same time, regardless of how many items were being tracked. Our results suggest that the capacity of multiple event monitoring reflects a limit on the number of items that can be monitored, but that within that small set of items it is possible to detect at least two events occurring at the same time.

Note that, even if it is possible to detect more than one event at a time, there may be a cost for that ability. In Experiment 2 we found that when there was only one event, the capacity was about 3.4, but the MEM capacity decreased to 2.7 when two events needed to be detected, though this conclusion depends on what one assumes about guessing. If guessing did not contribute to observers' performance in the two-target condition, then it could be argued that detection of two events was essentially the same as detection of a single event within the small set that can be monitored. Adding uncertainty about the timing of a second event, as was done in Experiment 3, reduced the size of the set that could be monitored (K = 2.3) compared to Experiment 2 (K = 2.7) where the two drop-off events always occurred at the same time (As tested with a mixed designed ANOVA with experiment type as a between-subjects factor: $F(1,23) = 6.78$, $p = 0.016$, $\eta_p^2 = 0.23$). This is in line with other studies that have also found that temporal uncertainty can modulate visual perception (Rolke & Hofmann, 2007; Westheimer & Ley, 1996).

In addition to temporal uncertainty, limits on visual working memory may also play a role in MEM. To successfully detect an event during tracking, observers must either notice the instant of change or they need to have registered the original state of the tracked items so that they could detect a change by comparison to the remembered prior state. Cohen, Pinto, Howe, and Horowitz (2011) found that the performance of tracking target location became worse when observers were also required to track the target identities. Thus, the difficulty of encoding and updating the status of the tracked items may also affect event tracking capacity. This may explain why we found a slightly higher capacity when observers monitored the simple items for an easier event (a circle dropping its bag in Exps 2 & 3) than when they monitored more complex items for a state change in Exp 1 and Wu and Wolfe (2016).

Interestingly, MEM capacity was identical across the different time intervals between the two changes (Experiment 1 & 3). One might have expected an attentional blink effect (Popple & Levi, 2007; Raymond, Shapiro, & Arnell, 1992) when the second target appeared 200–500 ms after the first. Nothing like an AB effect was found in our dual event detection tasks. The MEM performance was identical regardless of when the second event occurred, even when it fell within the typical time window of attentional blinks (Experiment 3).

Moreover, on blocks when observers knew that the events would occur asynchronously (Experiment 1), they might withhold their response after they detect the first target, waiting for the second event. If they miss the second event and the trial times out without a response, then the entire trial will be counted as an error and the observer would, in effect, lose credit for the one item that they found. On blocks when the events occur at the same time, observers would more likely respond as soon as they see any event. If need be, they could guess about the second event. This difference in strategy could explain why observers' performance was slightly better when both events occurred at the same time (Fig. 2). In an effort to assess these possible differences in strategy, we looked at the data from the blocks of Experiment 1 where the two events occurred at different times. On 41% of trials, only one target was successfully detect. Both targets were detected on 27% of trials and no target was detected on 32% of trials. The presence of 41% one-target trials shows that observers did not adopt a general strategy of allowing the trial to time out when they found only one target.

The results in the current study show that it is possible to detect two events simultaneously in the sustained monitoring task. This is consistent with reports of independent tracking mechanisms in multiple object tracking (Alvarez & Cavanagh, 2005) and multiple identity tracking (Hudson, Howe, & Little, 2012), though our experiments were not designed to specifically address this issue. Cavanagh and Alvarez (2005) showed that there are at least two, somewhat independent tracking mechanisms in each hemisphere. Their findings can be used to argue against the idea that there is always a single, undivided focus of attention. Our finding that two events can be detected at the same time is converging evidence for the possibility of multi-foci attention. If only one single focus of attention was available for the detection of two events, then performance should be worse when separate events occur at the same time. In addition, it is possible that the ability to monitor two events at the same time might require that those two events be similar to each other. Tracking performance might be impaired if the two events were different (e.g. monitor for one bag drop and one color change). Accounts that propose independent tracking mechanisms would seem to predict that observers could detect two different events (especially if they occurred in different visual fields/hemispheres), but a clear answer to this question would require further research.

In summary, our results show that the ability to monitor a dynamic scene for a discrete event is much more limited than the already limited ability to monitor the positions of moving items, and this capacity may be somewhat further constrained when a second event is added during tracking. However, it does appear that, within that very small set, it is possible to detect simultaneous events.

## Acknowledgment

The current research was funded by Army Research Office. No. R00000000000588.

## References

Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science, 16*(8), 637–643. http://dx.doi.org/10.1111/j.1467-9280.2005.01587.x.

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), http://dx.doi.org/10.1167/7.13.14 14.1–10.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325–14329. http://dx.doi.org/10.1073/pnas.0803390105.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436. http://dx.doi.org/10.1163/156856897X00357.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*(7), 349–354. http://dx.doi.org/10.1016/j.tics.2005.05.009.

Cohen, M. A., Pinto, Y., Howe, P. D. L., & Horowitz, T. S. (2011). The what–where trade-off in multiple-identity tracking. *Attention, Perception, & Psychophysics, 73*(5), 1422–1434. http://dx.doi.org/10.3758/s13414-011-0089-7.

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*(2), 172–184. http://dx.doi.org/10.3758/BF03193740.

Howe, P. D. L., Cohen, M. A., Pinto, Y., & Horowitz, T. S. (2010). Distinguishing between parallel and serial accounts of multiple object tracking. *Journal of Vision, 10*(8), 11. http://dx.doi.org/10.1167/9.8.239.

Hudson, C., Howe, P. D. L., & Little, D. R. (2012). Hemifield effects in multiple identity tracking. *PLoS One, 7*(8), http://dx.doi.org/10.1371/journal.pone.0043796.

Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in psychtoolbox-3? *Perception, 36*, S14. http://dx.doi.org/10.1068/v070821.

Makovski, T., & Jiang, Y. V. (2009). Feature binding in attentive tracking of distinct objects. *Visual Cognition, 17*(1–2), 180–194. http://dx.doi.org/10.1080/13506280802211334.

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*(5), 631–671. http://dx.doi.org/10.1080/13506280344000473.

Popple, A. V., & Levi, D. M. (2007). Attentional blinks as errors in temporal binding. *Vision Research, 47*(23), 2973–2981. http://dx.doi.org/10.1016/j.visres.2007.06.022.

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

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*. http://dx.doi.org/10.1037/0096-1523.18.3.849.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7*(1–3), 17–42. http://dx.doi.org/10.1080/135062800394667.

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychonomic Bulletin & Review, 14*(3), 522–526. http://dx.doi.org/10.3758/BF03194101.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261–267. http://dx.doi.org/10.1016/S1364-6613(97)01080-2.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16–20. http://dx.doi.org/10.1016/j.tics.2004.11.006.

Störmer, V. S., Winther, G. N., Li, S.-C., & Andersen, S. K. (2013). Sustained multifocal attentional enhancement of stimulus processing in early visual areas predicts tracking performance. *The Journal of Neuroscience, 33*(12), 5346–5351. http://dx.doi.org/10.1523/JNEUROSCI.4015-12.2013.

Westheimer, G., & Ley, E. (1996). Temporal uncertainty effects on orientation discrimination and stereoscopic thresholds. *Journal of the Optical Society of America, 13*, 884–886. http://dx.doi.org/10.1364/JOSAA.13.000884.

Wu, C.-C., & Wolfe, J. M. (2016). Multiple event monitoring. *Cognitive Research: Principles and Implications, 1*(1), 21. http://dx.doi.org/10.1186/s41235-016-0022-7.