ARTICLE IN PRESS

Vision Research xxx (xxxx) xxx–xxx

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

Please cite this article as: Wu, C.-C., Vision Research (2018), https://doi.org/10.1016/j.visres.2017.10.009

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