Attentional capture in multiple object tracking

Sebastian Pichlmeier
Institute of Psychology, Karlsruhe University of Education, Karlsruhe, Germany

Till Pfeiffer
Institute of Psychology, Karlsruhe University of Education, Karlsruhe, Germany

Attentional processes are generally assumed to be involved in multiple object tracking (MOT). The attentional capture paradigm is regularly used to study conditions of attentional control. It has up to now not been used to assess influences of sudden onset distractor stimuli in MOT. We investigated whether attentional capture does occur in MOT: Are onset distractors processed at all in dynamic attentional tasks? We found that sudden onset distractors were effective in lowering probe detection, thus demonstrating attentional capture. Tracking performance as dependent measure was not affected. The attentional capture effect persisted in conditions of higher tracking load (Experiment 2) and was dramatically increased in lower presentation frequency of the onset distractor (Experiment 3). Tracking performance was shown to suffer only when onset distractors were presented serially with very short time gaps in between, thus effectively disturbing re-engaging attention on the tracking set (Experiment 4). We discuss that rapid dis- and re-engagement of the attention process on target objects and an additional more basic process that continuously provides location information allow managing strong disruptions of attention during tracking.

Attentional capture in multiple object tracking

Research based on the multiple object tracking (MOT) paradigm (first described in Pylyshyn & Storm, 1988) has demonstrated that the human visual system is capable of keeping track of dynamically changing positions of multiple objects. In typical Experiments, observers follow a number of predefined targets out of a set of identical objects presented on a screen for an interval of time. Processes of visual attention are the central components in most current models of MOT (for reviews, see Cavanagh & Alvarez, 2005; Meyerhoff, Papenmeier, & Huff, 2017; Scholl, 2009). The MOT task ecologically resembles activities, like driving in heavy traffic, monitoring other players in team sports, or air planes on radar screens in air traffic control, and many more.

Concepts of visual attention generally include as central characteristics limited capacity and selection (Buschman & Kastner, 2015; Carrasco, 2011; Desimone & Duncan, 1995). Attentional selection of stimuli is assumed to be accomplished either endogenously or exogenously. Stimuli are selected because of voluntary, endogenous, modulation of attention or they are selected because events occurring in the world automatically, exogenously, draw attention toward these stimuli (Awh, Belopolsky, & Theeuwes, 2012; Jonides, 1981; Posner, Snyder, & Davison, 1980, additionally discuss selection history). Such latter involuntary shifts of attention are also referred to as attentional capture (for reviews, see Burnham, 2007; Rauschenberger, 2003; Ruz & Lupiañez, 2002; Theeuwes, 2010). Whether this attentional capture is modulated by top-down attentional control or whether it is strictly stimulus-driven is a debate still not resolved (Gaspelin, Ruthruff, & Lien, 2016; Luck, Gaspelin, Folk, Remington, & Theeuwes, 2020).

Processes of attention have been empirically demonstrated in object tracking (e.g. Iordanescu, Grabowecky, & Suzuki, 2009; Tombu & Seiffert, 2008; Tran & Hoffman, 2016). It is an interesting question whether these can be effectively distracted. Pylyshyn (2001) assumed core processes of MOT to be pre-attentive. According to Pylyshyn (2001, p. 146) the loss of an object during tracking could be due to an index being "grabbed" by the "appearance of a new visual object." But up to now there is no empirical evidence that a momentary interruption of pre-attentive or attentional processes is detrimental to MOT. This seems surprising as there is a plethora of studies for other classic paradigms of attention (as visual search or rapid serial visual presentation [RSVP]; for reviews see Rauschenberger, 2003; Ruz & Lupiañez, 2002; Theeuwes, 2010) investigating consequences of interruption of attention.

Citation: Pichlmeier, S., & Pfeiffer, T. (2021). Attentional capture in multiple object tracking. *Journal of Vision*, *21*(8):16, 1–20, https://doi.org/10.1167/jov.21.8.16.

https://doi.org/10.1167/jov.21.8.16 Received February 16, 2021; published August 11, 2021 ISSN 1534-7362

There is indeed some evidence from a series of studies that demonstrate enduring stability of MOT. Studies referred to as target recovery tasks (St. Clair, Huff, & Seiffert, 2010) have shown that tracking performance does not suffer when objects suddenly disappear for some time (Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005; Fencsik, Klieger, & Horowitz, 2007; Horowitz, Birnkrant, Fencsik, Tran, & Wolfe, 2006; Keane & Pylyshyn, 2006). Gap times in these experiments ranged from 107 to even 900 ms. In all of these studies, observers were able to compensate the temporary loss of online information and successfully continued tracking when objects reappeared. A tracking study by Bahrami (2003) closely resembles an attentional capture paradigm. He presented unexpected events (mud splashes) during tracking to hide changes within objects. Of interest here is the result that, although the mud splashes suddenly covering the entire screen could be expected to cause massive interruptions, they did not affect performance. Feria (2012) presented additional distractors with distinct features in an MOT-task and found no negative effect on performance in comparison to a no additional distractors condition; she discussed this finding in terms of a potential attentional capture effect that this stimulus constellation might have had in visual search but did not show in tracking.

Although, as these studies imply, MOT seems unaffected by sudden or salient events, strong bottom-up or stimulus-based influences have been shown to influence tracking nevertheless. Makovski and Jiang (2009), for example, found detrimental effects of target-distractor pairing; Erlikhman, Keane, Mettler, Horowitz, and Kellman (2013) comprehensively demonstrated an influence of certain stimulus features, although they were detrimental to performance and irrelevant to task goals. In their experiments, targets and distractors shared features leading to automatic grouping thus impairing performance. This poses an interesting problem. Most researchers in tracking would agree that tracking of objects is accomplished by applying attentional mechanisms. From attentional capture studies we know that most attentional mechanisms can be captured; an effect that has been shown to be stable even under more general conditions (e.g. Forster & Lavie, 2008a; Forster & Lavie, 2008b). It seems unlikely that attention in tracking is impermeable to additional distractors that onset abruptly and are irrelevant to the tracking task. Every dynamic action – walking, driving, carefully monitoring multiple streams of action in general – requires a capability to adapt to sudden unexpected changes. Therefore, it seems rather implausible that unexpected events are not registered when they happen. On the other side, it certainly would not be adaptive if dynamic actions broke down every time something unexpected happened.

Generally, attentional capture is operationalized by detrimental (but also beneficial in case the target is cued) influence of cues or irrelevant distractors on performance (Rauschenberger, 2003; Theeuwes, 2018) in paradigms like visual search or RSVP. If a cue or an irrelevant distractor is present in a trial, reaction time or error rate increase, thus indicating that attention has been captured. Following this conceptualization tracking performance would suffer when unexpected stimuli are displayed in MOT. In target recovery studies MOT was interrupted for periods lasting up to 900 ms (Keane & Pylyshyn, 2006). However, the temporal course of attentional capture is estimated to last only maximal 200 ms (Kim & Cave, 1999; Lamy, Tsal, & Egeth, 2003). Theeuwes, Atchley, and Kramer (2000) systematically varied the stimulus onset asynchronies (SOA) between presentation of distractor and search display. In these studies, attentional prioritization was first found on the salient irrelevant distractors and latest after 200 ms back again on the targets. These results are corroborated by a study using single cell recordings in primates trained to perform attentional tasks (Ogawa & Komatsu, 2004).

Considering these time courses, a plausible hypothesis would be that attentional capture does happen during MOT. The processes of capture of attention and subsequent disengagement and re-engagement on the relevant targets only happen too fast to affect tracking substantially (Gaspelin, Leonard, & Luck, 2017; Theeuwes, 2010; see also Gaspelin et al., 2016, for a similar concept of actual attentional capture with costs too small to be detected). Thus, a more time critical method is required to detect attentional capture in dynamic attentional tasks, such as MOT. Such a method would allow to directly measure allocation of attention on objects during MOT and thus to gain a deeper understanding of dynamic attention.

In order to measure effects of involuntary shifts of attention during MOT, we used the following experimental procedure: As we assumed the representation accompanying MOT to be very robust, very salient semantic stimuli (colored pictures of insects) were chosen as additional, sudden onset distractors assuming that such stimuli plausibly do attract attention. We follow here Forster and Lavie's approach to study the effects of irrelevant but salient distractors (Forster & Lavie, 2008a; Forster & Lavie, 2008b). Lavie and Forster used well-known cartoon figures as distractor stimuli because of their saliency and their definite irrelevance to the task at hand. Such semantic stimuli offer a promising method to provoke attentional capture in a MOT task.

We assessed attentional allocation using the probe dot technique, which allows high temporal resolution (Kim and Cave, 1999; Lamy et al., 2003; see also Posner, 1980; Posner et al., 1980) and has been used effectively in MOT studies (Alvarez & Scholl, 2005;

Flombaum, Scholl, & Pylyshyn, 2008; Huff, Papenmeier, & Zacks, 2012; Pylyshyn, 2006; Pylyshyn, Haladjian, King, & Reilly, 2008). Doran and Hoffman (2010) discuss probe detection as a means to assess the allocation of visual attention. They found that in some situations with relatively low tracking load (2 targets/ 2 distractors) electrophysiological measurements of probe registration showed no effect of probe location on the amplitudes of either anterior or posterior N1 component, whereas behavioral measures were significantly affected by this factor. However, with tracking loads higher than four objects electrophysiological data corresponded with behavioral data.

We collected behavioral data of detection of probes in tracking situations with relatively high load (4 targets/ 4 distractors and higher) and introduced a methodological variation: instead of using probes to detect the presence of attention on distractors, we were interested in detecting the absence of attention on targets by using number of missed probes as the central indicator. This was accomplished by dot-probing locations on which attention should reside in accordance with the current task goals - the target objects - but was expected to be dislocated by an additional onset distractor. The background for this is the still pending question of how attention is distributed in MOT. Prominent theories of MOT provide differing answers: Models that assume multiple foci (Cavanagh & Alvarez, 2005) or indexes (Pylyshyn & Storm, 1988); models that assume only one central attentional focus (Oksama & Hyönä, 2008); other models propose a distribution of attention among a group of objects (Yantis, 1992). Yet, other models assume a flexible resource (see Alvarez & Franconeri, 2007, for a discussion). The method used in our study complies with all these theoretical concepts. By focusing on absence of attention on targets we were able to rule out the possibility that only some of the attentional resource or that only one of many attentional foci or indexes would be captured. Thus, we were able to operationalize attentional capture as a loss of attention on an object in terms of missed probes - a definition that meets conditions of every model mentioned above.

We hypothesized that attention in MOT can be captured contrary to what target recovery studies imply (Alvarez et al., 2005; Fencsik et al., 2007; Horowitz et al., 2006; Keane & Pylyshyn, 2006). As Theeuwes et al. (2000) showed, the absence of performance cost in an attentional capture paradigm does not necessarily mean that attentional allocation is not influenced. If we assume that registering of distracting events does take place while not substantially harming performance itself, we are then dealing with a problem of sensitivity of experimental design. We claim that additional salient onset distractors do effectively grab attention.

To demonstrate this, we presented probes on targets following the additional onset distractors immediately – with a temporal distance of 100 ms – shorter than an attentional dis- and re-engagement would require. If these probes were detected less likely than probes appearing later in a trial, this could be taken as evidence of attentional capture in MOT. Additionally, we presented probes with a longer SOA (1500–2500 ms), which served the purpose to prevent the response strategy to react to onset distractors as signals of upcoming probes.

Overview of experiments

Experiment 1 demonstrated that attention was effectively distracted from the targets by sudden onset distractors. Experiment 2 provided evidence that capture effects are still detectable in MOT under higher task load. Experiment 3 showed that a lower frequency of onset distractors dramatically increased attentional capture. Experiment 4 finally demonstrated that the attentional capture effect can indeed be seen in task performance provided that disruptions by onset distractors are massive.

Experiment 1

Method

Participants

In our experiments, we were trying to measure an attentional capture effect conceptualized as dislocation of attention measured by a probe detection method. In pilot studies ($N = 27$) we found a substantial effect size: Cohen's $d > 1$ (55% probes detected less if probes were preceded by an onset distractor; R, R Core Team 2018, and the "effsize"-package by Torchiano, 2020, were used for estimates of Cohen's d). Additionally, we were also interested in tracking performance as a second dependent variable. We defined loss of half an object as a theoretically relevant measure for attentional capture. On the basis of a different pilot sample ($N = 72$), we determined an effect size of 0.9 (Cohen's d) to correspond with this definition, which equals a "large" effect (Cohen, 1992), which seems adequate in comparison to the effect size for the probe detection data. In order to be able to demonstrate effects of the latter dependent variable, the lower effect size was used as the basis for a sample size estimate. A power analysis based on an alpha value of .05 and beta of .95 was conducted. In the Results section, several comparisons and interactions are reported. The effect of interest, however, is the effect that the onset distractor has on probe detection and tracking performance, respectively. Therefore, we estimated sample sizes for comparisons

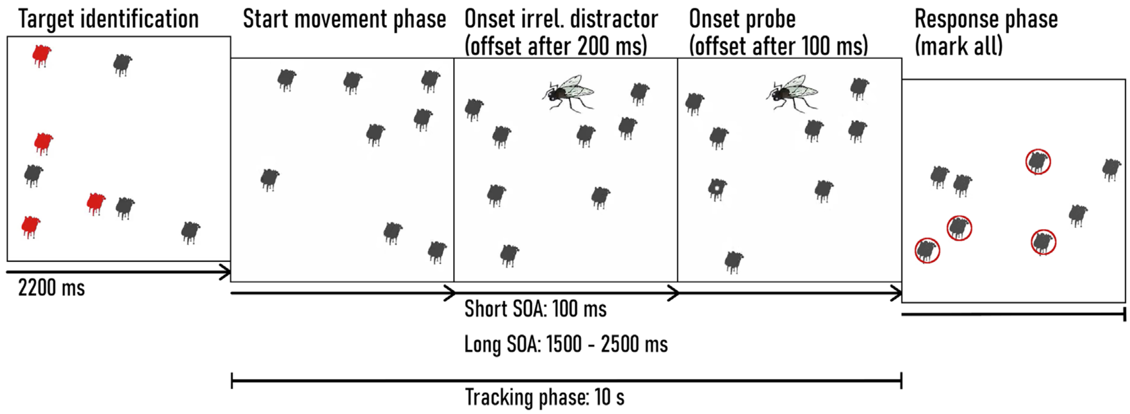

Figure 1. Illustration of the general procedure used in Experiments 1–3. Go to http://dx.doi.org/10.23668/psycharchives.4887 for example trials.

of two means. We found a minimal sample size of least 19 subjects was required in order to observe a possible effect of presentation of onset distractors on tracking ("pwr"-package; Champely, Ekstrom, Dalgaard, Gill, Weibelzahl, Anandkumar, Ford, Volcic, & De Rosario, 2020).

Twenty-two observers (12 women and 10 men) were included in Experiment 1. All had either normal or corrected-to-normal visual acuity. Participants ranged in age from 21.4 to 39.2 years ($M = 25.2$ years, $SD = 4.3$). All observers provided informed consent and the study was approved by the Human Ethics Advisory Group of the University of Education, Karlsruhe.

Apparatus

Stimuli were presented on a 21-inch monitor set to a resolution of 1920 by 1080 at a refresh rate of 60 Hz. The experiment was programmed in PsychoPy (Peirce, Gray, Simpson, MacAskill, Höchenberger, Sogo, Kastman, & Lindeløv, 2019). Participants were seated at a distance of 60 cm to the screen.

Stimuli

As tracking objects, we used irregular shapes with an approximate diameter 2 degrees of visual angle (see Figure 1) that vaguely resemble sheep (stimuli are also in use in parallel developmental studies). The objects were a uniform gray with a brightness value of 50% (hue, saturation, and value [HSV] = 0 0 50).

As probes, we presented amorphous white circles in the middle of the targets for 100 ms. These probes had a diameter of 0.33 degrees and had the color of the screen background.

We decided to use very salient visual stimuli as the additional onset distractors sensu Forster and Lavie (2008a; Forster and Lavie, 2008b) as we assumed a very robust representation for multiple moving objects. The mode of presentation of the additional distractor stimuli resembles that of onset singletons as used in the attentional capture methodology (for a discussion, see Liao & Yeh, 2013; Yantis, 2000). We refer to this kind of distractors as "onset distractors" in order to differentiate them from the regular MOT distractors. We used pictures of seven insects taken from Rossion and Pourtois (2004) that were at least 2.2 up to 3.4 times the size of the stimuli of the tracking set (measured using the largest extent in any direction). The onset distractors were shown for 200 ms.

Observers were informed that in some trial pictures of insects - the onset distractors - might appear and were instructed to ignore them.

Design and procedure

The experiment consisted of 60 trials for each observer. There were four or five probes per trial. Location of probes was randomized with the following restrictions: in half of the trials, the first probe was on a target. In order to avoid sequence effects, maximally two (4 probes per trial) or three probes (5 probes per trial) in any trial appeared sequentially on identical types of objects. Probes were presented randomly in the movement phase with a minimal interval of 1500 ms between two probe onsets, which was also the minimal interval for the first probe in the movement phase. The last probe was presented at least 1000 ms before the end of the movement phase. Keypresses of observers within 1000 ms following a probe were registered as a reaction to it. As location of within object type was not restricted, probes could appear on the exact same object within one trial. Probes were presented on distractors also in order to avoid that probes became informative of the role of the object they were presented upon. Our design had one factor with three levels: mode of presentation of onset distractor with the levels *no onset distractor* preceding the probe, *an onset distractor with a short SOA* (100 ms) preceding the probe, or *an onset distractor with a long SOA* (1500–2500 ms) preceding the probe. The factorial level of a long SOA was introduced in order to prevent a response strategy of reacting to onset distractors instead of probes. The temporal placement of an onset distractor was randomized with two exceptions: an onset distractor was never presented before the first probe in a trial, and the interval between the probes in which the onset distractor was presented had to be long enough to accommodate the respective SOA.

At the beginning of each trial, eight objects were presented on a white background (24.45 degrees × 24.45 degrees) in randomly selected positions. Four of the objects flashed red for six times over the course of 2.2 seconds thus identifying them as targets. All the objects then moved haphazardly through the display for 10 seconds. Each trial contained four or five probes. For each observer there were 30 trials without an onset distractor and 30 trials with an onset distractor.

Objects mostly moved in a fixed direction with a speed of 5.62 degrees/s, but on each frame, one randomly selected object had a 1% chance of turning. When an object did turn, the angle was selected randomly from between 1 degree and 359 degrees. If any two objects came within 1.82 degrees of one another, measured center to center, they were repelled. We controlled the number of collisions between objects of either role and surrounding square except for distractor/distractor collisions. Onset distractors were placed so they never touched targets, distractors, or surrounding square. Trials with more than 50 collisions in any category were replaced.

At the end of each trial, participants mouse-clicked all targets (mark-all). Each object could only be selected once, deselection was not possible. Objects which were selected were surrounded with a red circle (see Figure 1). After selection of four objects, participants received feedback about the number of correctly identified targets, about the number of correctly identified probes on targets, and the number of false alarms, which consisted of reactions to probes on distractors and of reactions without any preceding probe.

False alarm rates were registered in order to control for a systematic influence of onset distractors on false alarms in probe detection. False alarm rate was not included in probe detection analysis as the critical reactions in our study were reaction to probes with certain SOAs (short versus long) in relation to the onset distractor. In the sense of signal detection theory, there is no specific false alarm for these critical reactions. We expected a rather high false alarm rate in our trials, as several probes were presented per trial and subjects were instructed not to respond to probes on distractors.

Results

Data of this and the following Experiments 1 were analyzed in R (R Core Team, 2018) with the "ez"-package (Lawrence, 2016). To analyze the influence of the presentation of salient onset distractors on tracking of multiple objects, we examined two different dependent variables: first tracking performance in order to determine whether onset distractors harm tracking, and second, probe detection in order to determine whether onset distractors influenced allocation of attention.

Tracking performance – Experiment 1

Tracking performance in trials with onset distractors was not significantly worse compared with trials without onset distractors (t[21] = 0.87, $p = 0.392$, Cohen's $d = 0.19$, 95% CI [−0.42, 0.79]). In trials with onset distractors, objects were tracked marginally better with a numerical difference of 0.03 objects (see Table 1).

Probe detection – Experiment 1

In order to determine whether onset distractors attracted attention, we analyzed rates of probe detection. False alarm rate in Experiment 1 was fairly high (0.38/trial). Importantly, rate of false alarms was not influenced by presentation of the onset distractor (t [21] = 1.59, $p = .127$, Cohen's $d = 0.05$, 95% CI [−0.11, 0.01]) with less false alarms in trials with the onset distractor (0.37/trial) than in trials without (0.40/trial).

The mean rates of probe detection in this experiment were calculated as a function of modes of distractor presentation and then entered into a one-way within-subject analysis of variance (ANOVA), with the levels: no onset distractor, onset distractor with short SOA, and onset distractor with long SOA. Table 2 presents the mean results in the different experimental conditions (see also Figure 2). In contrast to tracking, performance probe detection was significantly influenced by modes of distractor presentation (F[1, 29] = 35.81, $p < 0.001$, $\eta^2_G = 0.276$; generalized Eta-Squared measure of effect size; Bakeman, 2005). The Greenhouse-Geisser adjustment was used to correct for violations of sphericity in this and the following analyses as necessary.

Planned comparisons of conditions using paired-samples t-tests were conducted in order to differentiate the effects of short versus long SOA. We expected a strong influence on probes with short SOAs and we expected probe detection to fully recover for probes

		Modes of presentation of onset distractor		
		No onset distractor	Onset distractor	Total
Exp. 1	Number of objects	3.65 (0.082)	3.68 (0.059)	3.66 (0.069)
	Percentage	91.21 (0.020)	92.01 (0.015)	91.61 (0.017)
Exp. 2	Number of objects	3.68 (0.153)	3.71 (0.145)	3.70 (0.148)
	Percentage	73.68 (0.031)	74.29 (0.029)	73.99 (0.030)
Exp. 3 Low frq	Number of objects	3.63 (0.062)	3.61 (0.063)	3.63 (0.060)
	Percentage	90.77 (0.016)	90.23 (0.016)	90.68 (0.015)
Exp. 3 High frq	Number of objects	3.68 (0.066)	3.70 (0.054)	3.69 (0.058)
	Percentage	92.08 (0.017)	92.50 (0.013)	92.29 (0.015)
Exp. 4	Number of objects	3.63 (0.113)	3.51 (0.115)	3.57 (0.113)
	Percentage	90.63 (0.028)	87.81 (0.029)	89.22 (0.028)

Table 1. Means of tracking performance in number of objects and in percent correctly identified in Experiment 1 (8 objects), Experiment 2 (10 objects), Experiment 3 (8 objects, low versus high frequency), and Experiment 4 (8 objects, high within trial frequency) for trials without an onset distractor, for trials with an onset distractor and for overall tracking performance (SE in parentheses).
No Onset Distractor = trials in which no onset distractor appeared; Onset Distractor = all trials with an onset distractor including trials with onset distractor that appeared 100 ms before presentation of a probe and trials with onset distractor that appeared between 1500 and 2500 ms before presentation of a probe; Total = Tracking Performance in all trials of an experiment (frq: frequency).

	Modes of presentation of onset distractor		
	Short SOA	Long SOA	Control
Exp. 1	70.30 (0.041)	88.79 (0.029)	89.39 (0.020)
Exp. 2	52.53 (0.049)	67.73 (0.047)	73.35 (0.040)
Exp. 3 hi frq	71.82 (0.031)	84.85 (0.033)	86.53 (0.025)
Exp. 3 low frq	38.18 (0.043)	83.64 (0.036)	89.36 (0.020)
Exp. 4	88.37 (0.018)		92.47 (0.012)

Table 2. Means of probe detection in Experiment 1 (8 objects), Experiment 2 (10 objects), Experiment 3 (8 objects, lower frequency), and Experiment 4 in percent for trials with onset distractors and for trials without an onset distractor (SE in parentheses).
Short SOA = trials with onset distractor that appeared 100 ms before presentation of a probe; Long SOA = trials with onset distractor that appeared between 1500 and 2500 ms before presentation of a probe; Control = trials in which no onset distractor appeared. The distinction short vs. long SOA does not apply to Experiment 4.

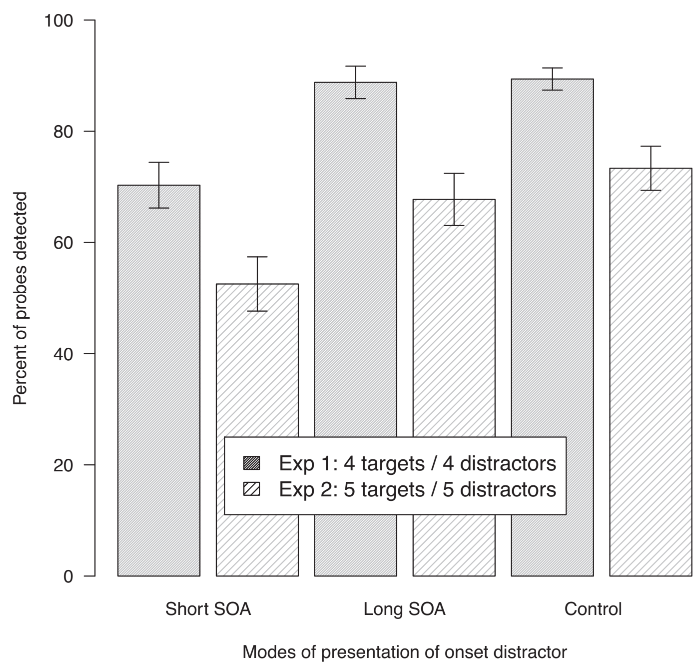

Figure 2. Effect of presentation of an onset distractor on means of probe detection as a function of task load (number of objects) for Experiment 1 (low load) and Experiment 2 (high load). Error bars represent Standard Error of Mean.

with long SOAs. The difference of probe detection between probes with short SOA and probes following no onset distractor was highly significant (t [21] = 6.21, $p < 0.001$, Cohen's $d = 1.32$, 95% CI [0.65, 1.99]). There was a marginal difference between the detection of probes with long SOAs and the detection of probes following no onset distractor in favor of the latter that did not reach significance (t [21] = 0.39, $p = .699$, Cohen's $d = 0.08$, 95% CI [−0.53, 0.69]).

Discussion of experiment 1

In Experiment 1, we found no influence of sudden presentations of very salient onset distractors on tracking, as was implied by tracking recovery studies (Alvarez et al., 2005; Fencsik et al., 2007; Horowitz et al., 2006; Keane & Pylyshyn, 2006). However, we were able to clearly show that attention was captured nevertheless by using a time-critical measure. Thus, our findings provide evidence that online allocation of attention in a dynamic task was effectively influenced by sudden presentation of salient onset distractors.

These findings seem to be in accordance with claims that MOT is managed by more than a single process. Assumptions of an additional process component have been discussed in various models: Doran and Hoffman's (2010) found electrophysiological results that the investment of attention is optional in MOT when tracking load is low. Fencsik et al. (2007) and Horowitz, Klieger, Fencsik, Yang, Alvarez, and Wolfe (2007) discuss a two-systems account that contains a position tracking mechanism and a focal attentional mechanism. Oksama and Hyönä (2008) propose a mixed model for tracking of identities that comprises a parallel and a serial attentional component. Störmer, Winther, Li, and Andersen (2013) assume a hybrid model consisting of an early visual process that allows parallel selection of multiple objects and at later processing stages a serial attentional mechanism. Pylyshyn (2001) explicitly does not preclude attention beside the preconceptual indexes to be involved in tracking. He considers attention among other functions to be necessary to reactivate fading object-index-bindings.

As tracking performance was unaffected and yet attention was captured in Experiment 1, the assumption that some component was active to serve as a beacon to direct attention back to the tracking set seems plausible. Such a beacon would have to be informative of the centroid of the targets at least (Huff, Papenmeier, Jahn, & Hesse, 2010).

The finding that the onset distractors left tracking unharmed is conceptually interesting. Strictly speaking, it could be argued that attentional capture can only be demonstrated if there is a performance loss (or gain) in the central task (Rauschenberger, 2003). Although we generally agree with this concept, we do follow Theeuwes et al. (2000) in the approach that sometimes an attentional capture is too fast to be recorded with conventional means (see also Gaspelin et al., 2016, for a similar concept). Theeuwes et al. (2000) demonstrated that - given an SOA longer than 150 ms - a singleton no longer interfered with search reaction time in their paradigm. To demonstrate that attention had been captured nevertheless, Theeuwes et al. showed that in trials with longer SOAs congruency effects still persisted showing that attention had indeed resided on distractors before it was quickly disengaged to return to the relevant targets.

We assume that a similar phenomenon is at work here in the tracking domain. As it becomes clear from our probe detection data, probes were detected almost 20% less when a salient distractor was presented very shortly before that probe. After 1500 to 2500 ms, however, this effect was no longer detectable. In summary, the pronounced reduction of about 20% in probe detection in trials with an onset distractor immediately preceding probes (SOA = 100 ms) shows that attention was distracted and thus effectively drawn from target objects.

Experiment 2: Load effects

How robust is this effect of attentional capture in MOT? Given the effect of the onset distractors in Experiment 1, we assumed that this effect would hold even under more demanding tracking conditions.

Lavie's perceptual load theory (Lavie & de Fockert, 2005; Lavie, Hirst, de Fockert, & Viding, 2004; Lavie & Tsal, 1994) implies that our finding could depend on load of the underlying task. It has been shown that task load can have pronounced effects on attention. Lavie and Fox (2000), for example, demonstrated that negative priming effects vanished completely when load was increased. In the tracking domain, Feria (2012) demonstrated that it depended on task load whether degree of distinctiveness of distractors enhanced tracking.

Effects of load on attentional capture have been shown in static attentional paradigms, like visual search. Forster and Lavie (2008b) found that attentional capture by salient cartoon figures was influenced by task load. In their experiments, they presented search displays containing six letters in the high and low load condition. The critical manipulation was that in the high load condition there were five different letters that shared features with the target letter whereas in the low load condition the distractor letters consisted of five o's.

Cosman and Vecera (2009) confirmed effects of load on attentional capture in visual search. In hybrid flanker and feature search tasks, they demonstrated that attentional capture depended on a low perceptual load setting. Stimulus sets in their experiments consisted either of one relevant letter plus two task-irrelevant flankers for low load conditions or of six relevant letters plus two task-irrelevant flankers for high load conditions. Irrelevant distractors in this series of experiments were either onset or offset flankers (Cosman & Vecera, 2009) or motion-onset flankers (Cosman & Vecera 2010a).

Can these findings from the visual search domain be generalized to a dynamic attentional task like MOT? A range of stimulus factors has been shown to influence tracking performance – some of them are speed, object size, set size, density, and crowding of objects (see, for example, Bettencourt & Somers, 2009; Cavanagh & Alvarez, 2005; Feria, 2012). These factors are potential candidates for load manipulation. We increased set size (from 8, 4 targets/4 distractors, to 10 objects, 5 targets/5 distractors) to assess the effect of load in MOT, as this manipulation has been successfully used in MOT (Feria, 2012) and it resembles that of studies of effects of load on attentional capture in visual search most.

Method

Participants

Twenty-five observers (16 women and 9 men) participated in the experiment (estimates of minimal sample size follow Experiment 1). All had either normal or corrected-to-normal visual acuity. Participants ranged in age from 18.2 to 53.3 years ($M = 24.7$ years, $SD = 6.5$). All observers provided informed consent and the study was approved by the Human Ethics Advisory Group of the University of Education, Karlsruhe.

Apparatus, stimuli, design, and procedure

Aside from the higher load with five targets among five distractors, apparatus, stimuli, design, and procedure were identical to Experiment 1.

Results

As in Experiment 1, we examined two different dependent variables: tracking performance and probe detection.

Tracking performance – Experiment 2

Increasing load from eight to 10 objects affected overall tracking performance. Observers' performance deteriorated from 92% in tracking four targets out of eight objects in Experiment 1 to 74% when tracking five targets out of 10 objects in Experiment 2 (see Table 1). To determine whether onset distractors harmed tracking in Experiment 2 under conditions of higher load, we compared means of trials with or without such distractors. As in Experiment 1, tracking performance in trials with onset distractors was not significantly worse compared with trials without onset distractors ($t[24] = 0.83$, $p = 0.413$, Cohen's $d = 0.17$, 95% CI [−0.40, 0.74]). Objects in trials with onset distractors were tracked marginally better with a numerical difference of 0.03 objects (see Table 1).

Tracking performance – Comparison experiment 1 versus experiment 2

A comparison between Experiment 1 and Experiment 2 showed that increasing load from eight to 10 objects affected overall tracking performance. Observers' performance deteriorated from 92% in tracking four targets out of eight objects in Experiment 1 to 74% when tracking five targets out of 10 objects in Experiment 2 (see Table 1). To analyze whether this difference was significant and whether increasing load interacted with conditions of presentation of onset distractors, we performed a two-factor mixed ANOVA with between factor experiment and within factor modes of presentation of onset distractor. This showed that increasing tracking load had a significant effect on tracking performance. The main effect of experiment became significant ($F[1, 45] = 24.68$, $p < 0.001$, $\eta_G^2 = 0.348$). The main effect of modes of distractor presentation did not reach significance ($F[1, 45] = 1.46$, $p = 0.234$, $\eta_G^2 < 0.001$); neither did the interaction between factors experiment and modes of distractor presentation become significant ($F[1, 45] = 0.02$, $p = 0.876$, $\eta_G^2 < 0.001$).[2]

Probe detection – Experiment 2

Number of false alarm rate in Experiment 2 again was fairly high (0.87/trial). Importantly - as in Experiment 1 – rate of false alarms was not influenced by presentation of the onset distractor (t [24] = 1.30, $p = .205$, Cohen's $d = 0.07$, 95% CI [−0.04, 0.18]) with more false alarms in trials with the onset distractor (0.90/trial) than in trials without (0.84/trial).

As in Experiment 1, mean rates of probe detection were calculated as a function of modes of distractor presentation (see Table 2, Figure 2). A one-way within-subject ANOVA, with the levels: onset distractor with short SOA, onset distractor with long SOA, and no onset distractor showed that probe detection was significantly influenced by presentations of onset distractors ($F[2, 48] = 21.88$, $p < 0.001$, $\eta_G^2 = 0.136$).

Paired-samples t-tests were conducted as planned comparisons in order to differentiate the effects of short versus long SOA. As in Experiment 1, the difference in probe detection between probes with short SOA and probes following no onset distractor was highly significant (t [24] = 5.93, $p < .001$, Cohen's $d = 1.19$, 95% CI [0.57, 1.80]). The difference between the detection of probes with long SOAs and the detection of probes following no onset distractor in favor of the latter reached significance (t [24] = 2.37, $p = 0.026$, Cohen's $d = 0.47$, 95% CI [0.10, 1.05]) in Experiment 2.

Effect of load on attentional capture – Comparison experiment 1 versus experiment 2

Increasing load from eight to 10 objects also affected overall probe detection performance. Observers' performance deteriorated from 83% in Experiment 1 to 65% in Experiment 2. A two-factor mixed ANOVA with between factor Experiment and within factor modes of presentation of onset distractor showed that this deterioration did not interact with the factor of modes of distractor presentation.

The main effect of experiment became significant ($F[1, 45] = 12.86$, $p < 0.001$, $\eta_G^2 = 0.189$); as did the main effect of modes of distractor presentation ($F[2, 71] = 51.85$, $p < 0.001$, $\eta_G^2 = 0.177$). The interaction between factor experiment and modes of distractor presentation was not significant ($F[2, 71] = 0.72$, $p = 0.456$, $\eta_G^2 = 0.003$, see footnote 2). We attribute this finding of a nonsignificant interaction to the diminished power of the mixed-effect analysis with unequal sample sizes. Note that in Experiment 2 probes were detected significantly less in comparison to the control condition, whereas in Experiment 1 this difference did not reach significance.

Discussion of experiment 2

In summary, Experiment 2 confirmed the findings of Experiment 1 under high load conditions. As in Experiment 1, tracking performance was not influenced by the presentation of onset distractors, thus showing no signs of attentional capture. Task load did influence tracking performance substantially. Increasing load led to a substantial loss of 18%. However, in contrast to possible predictions based on load theory, probe detection still distinctly suffered even under the high load condition of Experiment 2 when probes appeared shortly (100 ms) after a salient onset distractor. In Experiment 2, the attentional capture effect even appeared prolonged. It was still detectable in the long SOA condition, which was added for methodological reasons - an unpredicted effect and thus a serendipitous result most likely due to the higher load demands of Experiment 2. Experiment 2 thus corroborates the findings of Experiment 1 and demonstrates the robustness of the attentional capture effect in MOT. Although general probe detection deteriorated significantly in comparison to Experiment 1, the attentional capture effect was not modulated by this decrease.

A modulation of attentional capture by task load has been described in various studies of visual search (e.g. Cosman & Vecera, 2009; Cosman & Vecera 2010a; Forster & Lavie, 2008b). For our study, we can definitely rule out that the extent of load manipulation was too small. Tracking performance suffered greatly as number of objects was raised - with a performance loss of 18% it was rather substantial. Thus, task load definitely was influenced. In contrast to findings of Cosman and Vecera (2009), however, we still found strong attentional capture in probe detection after load had been increased.

Why does load manipulation leave the attentional dislocation in our paradigm unaffected whereas it strongly influences it in visual search? One possible explanation could be based on Yantis and Jonides (1990) considerations on the automaticity of an attentional task. The current experiments seem to fulfill their load-insensitivity criterion. This could be taken as an indication that the irrelevant onset distractors in our paradigm capture attention in an automatic fashion.

Experiment 3: Effects of distractor frequency

Experiments 1 and 2 show the robustness of the attentional capture effect in our paradigm. The extent of the effect seems remarkable as probes are detected almost 20% less in Experiment 1 when they follow an onset distractor. However substantial this effect may be, in 70% of the trials observers did not miss the probes. Which factors are responsible that in spite of the considerable effect of the onset distractors partial cognitive control seems to remain? What factors do influence the extent of attentional capture?

The load manipulation of Experiment 2 only influenced tracking and overall probe detection rate. This manipulation, however, did not show any effect on the attentional capture effect. In order to achieve this, it seems promising to try to tackle the conditions of cue presentation itself more directly. One way that has proven effective in influencing attentional capture is to vary the frequency of the cues or irrelevant distractors in a task (Cosman & Vecera, 2010b; Forster & Lavie, 2008a; Neo & Chua, 2006; see Rauschenberger, 2003, for a discussion of influence of practice on the ability of an onset cue to capture attention).

Neo and Chua (2006) hypothesized that low frequency cues elicit orienting responses. High frequency presentation of cues supports the forming of a representation that incorporates these cues. Thus, cues in a high frequency condition can become part of this representation and no longer elicit orienting responses.

In order to test whether manipulation of cue frequency also influences allocation of attention in tracking, we added frequency as factor in Experiment 3. Because frequency of cues was already relatively high in Experiments 1 and 2, we compared this high frequency with a lower frequency condition in Experiment 3. Following Neo and Chua, we hypothesized that lowering the frequency of presentation of irrelevant onset distractors would result in an even more pronounced effect on probe detection thus increasing the attentional capture effect.

Method

Participants

Twenty-two observers (11 women and 11 men) participated in the experiment. All had either normal or corrected-to-normal visual acuity. Participants ranged in age from 18.0 to 56.5 years ($M = 24.5$ years, $SD = 7.8$). All observers provided informed consent and the study was approved by the Human Ethics Advisory Group of the University of Education, Karlsruhe.

Apparatus, stimuli, design, and procedure

The high frequency condition of Experiment 3 was identical to the design of Experiment 1 (15 distractors with short SOA, 15 distractors with long SOA in a total of 60 trials), the low frequency condition of Experiment 3 resembled Experiment 1 with the exception of a lower frequency of presentation of the onset distractor (5 distractors with short SOA, 5 distractors with long SOA in a total of 60 trials). We set a comparatively agreeable low frequency in order to keep the number of total trials within limits of observers' perseverance and in approximate accordance with preceding studies (Cosman & Vecera, 2010b: 20% vs. 80%; Forster and Lavie, 2008a: 10% vs. 50%; Neo & Chua, 2006: 18.75% vs. 75%). Both blocks were presented in a within-subjects design. As the high frequency block of Experiment 3 was a direct replication of Experiment 1, it followed the low-frequency block; thus, permitting to control for potential carry-over effects.

Results

As in the previous experiments, we based our analyses on two different dependent variables: tracking performance (see Table 1) and probe detection (see Table 2). In Experiment 3, we examined effects of frequency of presentation of an onset distractor on attentional capture effects.

Tracking performance – Experiment 3

We analyzed the effects of frequency of the onset distractor on tracking performance in Experiment 3 with a two-way within-subject ANOVA with the factors frequency and modes of distractor presentation. It showed that - as in Experiments 1 and 2 - the main effect of modes of distractor presentation did not become significant ($F[1, 21] < 1$, $p = 0.932$, $\eta_G^2 < 0.001$). However, the main effect of frequency of distractor presentation became significant, indicating that tracking was worse for the low frequency block ($F[1, 21] = 7.83$, $p = 0.011$, $\eta_G^2 = 0.016$). The interaction between these two factors was not significant ($F[1, 21] < 1$, $p = 0.507$, $\eta_G^2 = 0.001$).

Tracking performance – Comparison experiment 1 versus experiment 3 (high frequency condition)

The high frequency condition of Experiment 3 is a direct replication of Experiment 1. A possibly effective difference to Experiment 1 might be that the high frequency condition of Experiment 3 was presented as the second block in Experiment 3 following the low frequency block, which might have caused training or carry-over effects.

Overall tracking performance between Experiment 1 (91.6%) and the high frequency condition of Experiment 3 (92.3%) differed slightly (see Table 1). To analyze whether effects of position in a blocked presentation influenced tracking or interacted with modes of presentation of onset distractors, we performed a two-factor mixed ANOVA with between factor experiment and within factor modes of presentation of onset distractor. This showed that position had no significant effect on tracking performance. The main effect of experiment did not become significant ($F[1, 42] = 0.09$, $p = 0.764$, $\eta_G^2 = 0.002$). The main effect of modes of distractor presentation did also not reach significance ($F[1, 42] = 0.98$, $p = 0.327$, $\eta_G^2 = 0.001$); neither did the interaction between factors experiment and modes of distractor presentation become significant ($F[1, 42] = 0.10$, $p = 0.758$, $\eta_G^2 < 0.001$).

Probe detection – Experiment 3

Number of false alarm rate in Experiment 3 was relatively small (0.17/trial). Importantly - as in Experiments 1 and 2 – rate of false alarms was not influenced by presentation of the onset distractor ($F[1, 21] = 1.97$, $p = 0.175$, $\eta_G^2 = 0.009$) with less false alarms in trials with the onset distractor (0.15/trial) than in trials without (0.18/trial) - but it was influenced by frequency of presentation of onset distractor ($F[1, 21] = 5.03$, $p = 0.036$, $\eta_G^2 = 0.026$) with less false alarms in trials with frequent presentation of the onset distractor (0.14/trial versus 0.20/trial). These two factor did not interact ($F[1, 21] = 0.36$, $p = 0.557$, $\eta_G^2 = 0.002$).

Mean rates of probe detection were calculated as a function of modes of distractor presentation (see Table 2, Figure 3) and frequency. A two-way within-subject ANOVA showed that probe detection was significantly influenced by modes of presentations of onset distractors ($F[2, 33] = 99.62$, $p < 0.001$, $\eta_G^2 = 0.499$), frequency ($F[1, 21] = 29.37$, $p < 0.001$, $\eta_G^2 = 0.116$), and the interaction of these two factors ($F[1, 28] = 28.98$, $p < 0.001$, $\eta_G^2 = 0.234$).

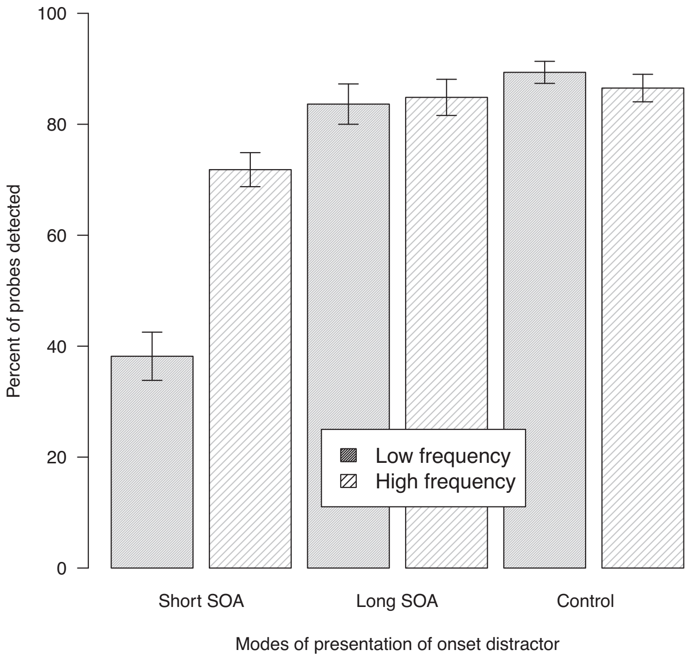

Figure 3. Effect of presentation of an onset distractor on means of probe detection as a function of frequency for Experiment 3. Error bars represent Standard Error of Mean.

Because of the significant interaction separate analyses were conducted for the two frequency blocks. The factor modes of distractor presentation became significant in the low frequency block (F[1, 28] = 79.73, $p < 0.001$, $\eta^2_G = 0.674$) and also in the high frequency block (F[2, 32] = 18.15, $p < 0.001$, $\eta^2_G = 0.191$).

Paired-samples t-tests were conducted as planned comparisons in order to differentiate the effects of short versus long SOA. Low frequency condition: the difference in probe detection between probes with short SOA and probes following no onset distractor was highly significant (t [21] = 13.09, $p < 0.001$, Cohen's d = 2.96, 95% CI [1.90, 4.02]). The difference between the detection of probes with long SOAs and the detection of probes following no onset distractor in favor of the latter did not reach significance (t [21] = 1.75, p = 0.094, Cohen's d = 0.39, 95% CI [−0.08, 0.86]). A similar pattern showed for the high frequency block: short SOA versus no onset distractor (t [21] = 4.85, $p <$ 0.001, Cohen's d = 1.16, 95% CI [0.52, 1.71]); long SOA versus no onset distractor (t [21] = 0.96, p = 0.350, Cohen's d = 0.11, 95% CI [−0.35, 0.16]).

Comparisons between low and high frequency blocks showed that frequency significantly affected the difference between rate of detection of probes that were presented shortly after an onset distractor (t [21] = 6.64, $p < 0.001$, Cohen's d = 1.90, 95% CI [0.93, 2.87]) with lower probe detection in the low frequency block; whereas detection rate of probes presented after a long SOA did not differ between low and high frequency blocks (t [21] = 0.34, p = 0.738, Cohen's d = 0.07, 95% CI [−0.37, 0.52]); however, the detection rate of probes that followed no onset distractor in the low frequency condition was significantly higher than in the high frequency block (t [21] = 3.35, p = 0.003, Cohen's d = 0.22, 95% CI [−0.09, 0.36]).

Probe detection – Comparison experiment 1 versus experiment 3 (high frequency condition)

As discussed above, the high frequency condition of Experiment 3 is a direct replication of Experiment 1 with the exception that the high frequency condition of Experiment 3 is the second part of a two block design. It is possible that probe detection performance might be influenced by position effects. Overall detection rate differed slightly: 82.8% in Experiment 1 versus 81.1% in the high frequency condition of Experiment 3 (for specific values, see Table 2).

A two-factor mixed ANOVA with between factor experiment and within factor modes of presentation of onset distractor showed that the factor experiment did not significantly influence probe detection rate (F[1, 42] = 0.22, p = 0.639, $\eta^2_G = 0.004$) nor did it interact with the factor modes of presentation of the onset distractor (F[1, 61] = 1.28, p = 0.292, $\eta^2_G = 0.007$).

As the factor experiment did not significantly affect probe detection, results for the main effect of modes of onset distractor presentation (F[1, 61] = 52.02, $p < 0.001$, $\eta^2_G = 0.234$) correspond with those of the separate analyses of Experiment 1 and the high frequency condition of Experiment 3.

Discussion of experiment 3

Probe detection

Experiment 3 replicated the general findings of the previous experiments. Again, the presentation of an additional onset distractor shortly before the appearance of a probe lowered detection probability of this probe. The high frequency condition of Experiment 3 also directly replicated the specific findings of Experiment 1, which showed that our findings were robust against sequence effects. However, presenting the onset distractor with a lower frequency dramatically increased the attentional capture effect in Experiment 3.

Our results thus parallel those of Forster and Lavie (2008a). In the study of Cosman and Vecera (2010b), the attentional capture effect vanished under conditions of high frequency distractor presentation and high load. Neo and Chua (2006) demonstrated the role of frequency as rare onset-cues still elicited an attentional capture even when targets were precued with 100% validity. They, however, found no attentional capture when onset-cues appeared more often.

Attentional capture was markedly affected by manipulation of frequency of distractor presentation in our experiment. Less than 40% of the probes that followed onset distractors after a 100 ms SOA in the low frequency condition were detected - a substantial difference to the more than 70% detection rate of the high frequency condition. Although probes after onset distractors with a short SOA were detected more often in the high frequency condition than in the low frequency condition, they nevertheless were detected significantly less in comparison to the control condition. Even though observers in the high frequency conditions knew well what to expect and were reliably and regularly confronted with onset stimuli besides the tracking set, they nevertheless could not effectively guard themselves against being distracted.

Tracking performance

In accordance with Experiments 1 and 2 there was no detectable effect of onset distractor presentation on tracking performance in Experiment 3. There was no difference between trials with or without an onset distractor. However, there was a significant effect of frequency of distractor presentation on tracking performance. General tracking performance was better in the block with high frequency presentation of onset distractors in comparison to the low frequency block. It seems not very likely that this performance advantage is due to a practice effect. Although the high frequency block followed the low frequency block in Experiment 3, it was a direct replication of Experiment 1 and the results of Experiment 1 do not significantly differ from those of the high frequency block of Experiment 3. A likely explanation for the higher tracking performance in the high frequency condition, is that because of the more challenging task situation more attentional resources are deployed to the tracking task.

Note, however, that a general change in task performance is not informative of a potential attentional capture effect on tracking performance. To demonstrate this, onset distractors have to be effective in reducing tracking performance in comparison to trials without such disruptions. Experiment 4 was designed to show that onset distractors actually do have the potential to achieve this.

Experiment 4: Disruption of performance

Experiments 1, 2, and 3 underlined the robustness of the tracking system against sudden disruptions. Although the onset distractors that were presented definitely were very salient, they nonetheless failed to harm tracking performance. We were able to demonstrate that attention was indeed captured and drawn toward the disrupting stimulus by applying a probe detection technique. Experiment 3 even demonstrated a dramatic loss of more than 50% of probes detected less in trials with disruption in comparison to control trials. However, we also found that only after a 1500–2500 ms probe, detection was reconstituted as shown in the unaffected detection rate in the long SOA condition. A small influence of the onset distractor presented with the long SOA interval could only be seen in the high load tracking situation of Experiment 2.

Therefore, it takes only a very short time after a disrupting event to re-establish attentional control of the tracking set thus maintaining unharmed tracking. Experiment 4 was designed to disturb this reconstitution phase by presenting yet another onset distractor while the attentional focus is redirected back to the tracking set. We hypothesized that provided that the process of re-establishing attentional control is constantly disturbed, more attentional resources have to be devoted to the process of reconstituting the tracking process. Therefore, we hypothesized that a high frequency presentation of onset distractors can effectively disrupt tracking.

Method

Participants

Sixteen observers (9 women and 7 men) participated in the experiment. Sample sizes were planned in analogy to the previous experiments. Due to developments of the COVID-19 pandemic, data collection was stopped prematurely. All had either normal or corrected-to-normal visual acuity. Participants ranged in age from 18.0 to 28.7 years ($M = 24.4$ years, $SD = 2.4$). All observers provided informed consent and the study was approved by the Human Ethics Advisory Group of the University of Education, Karlsruhe.

Apparatus, stimuli, design, and procedure

Experimental design resembled Experiment 1 with the exception of a manipulation of frequency of onset distractors within the experiment. In 30 trials, there were no onset distractors. In another 30 trials, 26 or 27 onset distractors appeared in the first 9 seconds of the tracking episode, which resulted in a mean ISI (offset last onset distractor to onset next onset distractor) of 13.3 ms (range = 8.3 ms to 18.3 ms). Thus, in critical trials, there were three to four onset distractors per second; in the control trials, there were no onset distractors at all. The resulting 60 trials were presented in randomized order. The scheme for probe presentation exactly resembled Experiment 1. The timing of presentation of onset distractors and of probes was independent in Experiment 4.

Results

As in the previous experiments, we based our analyses on two different dependent variables: tracking performance and probe detection. In Experiment 4, we examined the effects of a massive increase of presentations of an onset distractor within one trial on attentional capture effects.

Tracking performance – Experiment 4

In Experiment 4, frequency of presentation of onset distractors within a trial was increased. In contrast to the previous experiments, presentation of onset distractors significantly reduced tracking performance in comparison to trials without an onset distractor ($t[15] = 3.48$, $p = 0.003$, Cohen's $d = 0.25$, 95% CI [0.10, 0.39]; see Table 1).

Probe detection – Experiment 4

Number of false alarm rate in Experiment 4 was similar to that of Experiment 3 (0.23/trial). Importantly - as in Experiment 1 – rate of false alarms was not influenced by presentation of the onset distractor (($t\ [15] = 1.72$, $p = 0.106$, Cohen's $d = 0.40$, 95% CI [−0.09, 0.88]) with more false alarms in trials with the onset distractor (0.28/trial) than in trials without (0.17/trial).

As for the previous experiments, mean rates of probe detection were calculated as a function of modes of distractor presentation. In contrast to the previous experiments, however, there were no two conditions of distractor presentation. As there were about three or four onset distractor per second in a critical trial, individual SOAs were not distinguishable. Thus, probe detection in critical trials in general was compared with probe detection in trials without any onset distractor. As in Experiments 1 to 3, this difference became significant ($t[15] = 3.00$, $p = 0.009$, Cohen's $d = 0.62$, 95% CI [0.16, 1.08]), albeit the difference in this experiment was quite small compared to the effects in the previous experiments (see Table 2).

Discussion of experiment 4

As in the previous experiments, probe detection depended significantly on modes of distractor presentation. In contrast to previous situations, the attentional capture effect was much less pronounced in Experiment 4, whereas probe detection in control trials, however, was comparable to probe detection rates of control conditions of the previous experiments. A reduction of the effect might be due to a reduced coupling of onset distractors and probes such that there was no short SOA condition and possibly due to observers utilizing an adaptive strategy to cope with situations of continuous disruptions.

The central finding of Experiment 4 was that, in this experiment, tracking performance deteriorated for trials with onset distractors in comparison to control trials. Presentation of up to four onset distractors per second lead to effective disruption of the tracking process. Huff et al. (2010) studied the processes involved in critical tracking situations. They examined gaze behavior in tracking after substantial disruptions and found less gaze on targets only for the first 500 ms after a viewpoint change, a sudden change of the perspective of the observer on the tracking array, whereas gaze on the centroid of the targets remained unaffected. The authors assumed that looking at the centroid aided the process of re-establishing the tracking set. In Experiment 4, frequency of distraction of attention was so high that the interval needed to re-establish the tracking mechanism after a disrupting event was itself disrupted. As the experiments of Huff et al. (2010) imply, a specific mechanism serves to reconstitute tracking. The frequent presentation of the onset distractor presumably prevented this mechanism to become effective. The deterioration of tracking performance induced by this continuous dislocation of attention provides evidence that although an additional component is necessary to direct attention as a beacon, it does not suffice alone to maintain tracking. Doran and Hoffman (2010) provided electrophysiological results that showed that attentional mechanisms are only optional for low tracking loads well below the load used in our experiments.

One possible explanation for the decrease in tracking performance in Experiment 4 might be increased crowding (Bettencourt & Somers, 2009; Franconeri, Jonathan, & Scimeca, 2010; Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008; Scimeca & Franconeri, 2015). However, the number of objects in each single moment of a trial in Experiment 4 - eight objects plus one onset distractor - was less than in the high load condition of Experiment 2, in which there were 10 objects plus the occasional onset distractor. Mind that there was no moment in a trial in Experiment 4 in which there was more than one onset distractor. Thus, crowding as a "failure of individuation due to proximity" (Franconeri et al., 2008, p. 802) would have been more likely in Experiment 2, which showed no effect of onset distractor presentation on tracking performance. One further argument is worth considering when discussing crowding. Bae and

Flombaum (2012) showed that tracking performance was higher in trials in which nontargets changed to a different color from the color of the targets whenever a nontarget was within a critical distance of 4 degrees of visual angle from a target and thus showed that color changes can prevent tracking errors. In our experiment, the onset distractor was specifically chosen to be very salient and thus very different from the objects of the tracking set. Following Bae and Flombaum, the saliency of the onset distractors in our experiments would facilitate individuation and prevent tracking errors.

Temporal crowding, visual masking, or the attentional blink describe phenomena that are similar to the findings in our experiments. In each of these paradigms, the performance in processing of a stimulus is decreased when another stimulus precedes it (or follows it in the case of backward masking) within relatively short intervals: typically less than ±150 ms for masking (Enns & Di Lollo, 2000), typically more than 150 ms for temporal crowding (Yeshurun, Rashal, & Tkacz-Domb, 2015), and typically 200–500 ms for attentional blink (Dux & Marois, 2009). These paradigms, however, seem different to our findings as a central requirement for each of these is that the crucial stimuli are presented in "spatiotemporal vicinity" (Agaoglu, Breitmeyer, & Ogmen, 2018, p. 1). In masking and temporal crowding, critical stimuli are placed on the same position or are closely adjacent (Enns & di Lollo, 2000, Yeshurun et al., 2015); the attentional blink is typically demonstrated in a RSVP, where stimuli are typically placed in one RSVP-stream (Dux & Marois, 2009; but see Potter, Staub, & O'Connor, 2002, who used two streams separated by only 0.4 degrees).

In summary, Experiment 4 demonstrated an attentional capture effect resulting in reduced tracking performance and provided a potential explanation for the failure to provoke tracking performance disruptions in the previous experiments: although the onset distractors used were very salient and although they effectively captured attention tracking performance was unaffected, because the reconstitution of the attentional system takes only a very short time. Cavanagh, Holcombe, and Chou (2008) found an interval of even as little as 50 ms to be enough for dynamic attention to sufficiently process one stimulus before arrival of the next stimulus. Even if we account for some time of travel of attention from the stimulus of interest, a target, for example, to the disrupting stimulus and back, the processes of disruption and redeployment will not take long enough for the stimulus array to change enough that resuming tracking would become impossible. Because of this potential of dynamic attention to deal promptly with disruptions, monitoring dynamic scenes can be accomplished with high stability.

General discussion

We conducted four experiments to examine how visual attention is deployed in attentional tracking of dynamic scenes. Processes of visual attention are considered the central components in most theories of MOT. However, theories still differ on the exact form of these processes. Most concepts include either a number of tracking mechanisms, examples of such concepts are FLEXes – a flexible number of tracking mechanisms limited by a shared resource (Alvarez & Franconeri, 2007), or multifocal attention (Cavanagh & Alvarez, 2005) or - the preconceptual - sticky indexes of the FINST-theory (Pylyshyn, 2001); other models propose a single process that organizes objects into a virtual polygon (Yantis, 1992; see also Zhao, Gao, Ye, Zhou, Shui, & Shen, 2014, for evidence in multiple identity tracking) or a single attentional spotlight refreshing visual short term memory (Oksama & Hyönä, 2008) or object files (Kahneman, Treisman, & Gibbs, 1992). All these models share the common notion that some kind of visual attention (or preconceptual index) is deployed to the task at hand or - as in MOT - to various objects of interest. We were interested in the nature of this deployment and tried to influence it directly by trying to divert it from the focused objects. The current experiments demonstrated that it is indeed possible to capture attention set on multiple moving objects. This seems remarkable given the impressive robustness of MOT against interruptions (see also target recovery studies). We replicated our findings and specified some of the conditions that determine its extent. We were able to show that the attentional capture phenomenon we found was not diminished by task load; it appeared even prolonged in conditions of higher load - a provisional finding as it was not hypothesized a priori (Experiment 2). The attentional capture effect was increased substantially when frequency of onset distractors was lowered (Experiment 3). Finally, in Experiment 4, we demonstrated attentional capture in terms of deterioration of tracking performance.

Generally, attentional capture is operationalized by extent of performance change (cost or benefit) in trials with a cue (Rauschenberger, 2003). Although the onset-cues used in our experiments were chosen to be phenomenologically impressive and at best startling, there was no detrimental effect on tracking when these cues were presented on arbitrary points of time (Experiments 1, 2, and 3). As Kim and Cave (1999) and Lamy et al. (2003), we found attentional capture effects by probing for attention. Thus, we were able to show that immediately after a salient onset distractor appears during a tracking trial, attention moves rapidly away from the targets - and probably toward this new object. Huff et al. (2012) found diminished probe detection on the boundaries between two meaningful events. To register the interruption in the chain of events, their observers had to track the development of a basic story (a soccer match). In contrast to their study, we assume that the attentional dislocation described here is a process of involuntary orienting. Thus, a substantial amount of the probes following the salient onset distractors remained undetected, whereas other probes not preceded by onset distractors were registered in a rate nearly touching ceiling.

In contrast to Kim and Cave (1999) or Lamy et al. (2003), we did not measure attentional allocation by testing for the presence of attention. By means of salient cues we elicited a detachment of attention that led to a measurable absence of attention. The dependent measure in paradigms probing attention generally is reaction time. The underlying logic hereby is that if attention is captured, it takes longer to finally detect a probe resting on a target. What this method actually does is to sum up the course of attentional deployment in a given interval of time. It provides reliable and valid information about the distribution of attention. However, this measure does neither tell us anything about how effectively attention was captured nor about the exact temporal course of its allocation. The dynamic nature of the MOT task allowed us to use absence as an indirect measure. It can signal if attention resides on a certain location and it can do this with very high temporal resolution. On the grounds of this logic, we found that presenting a salient onset distractor led to a pronounced dislocation of attention. Probes that were undisturbed were detected in a rate of about 90% in Experiments 1 and 3. When salient distractors preceded probes by a very short interval of time, detection rate dropped markedly. By concentrating on the misses in probe detection, we were able to measure the time course of the dislocation of attention. With a stimulus onset asynchrony of 100 ms between onset distractor and probe, probe detection rate sank to a minimum of 38%. When onset distractors preceded probes with an SOA of 1500 to 2500 ms, probe detection was again fully reconstituted if task load was relatively low. In conditions with higher load (Experiment 2), the attentional capture effect was still detectable for probes with a longer SOA.

Our experiments provide a model explanation for the robustness of the MOT task against interruptions. It is not the case that attention is rigidly fixed on objects. It can be “grabbed” (Pylyshyn, 2001, p. 146) and dragged to other locations. However, attention is rapidly disengaged (see also Theeuwes, 2010 and Theeuwes, De Vries, & Godijn, 2003, for a thorough discussion of the temporal course of attentional disengagement) and returned to the relevant targets. In our experiments, attentional mechanisms returned back to the targets efficiently to take up tracking without any performance cost. As the target recovery studies (see above) already demonstrated, it is possible to take up tracking without any cost after interruptions of up to 900 ms (Keane & Pylyshyn, 2006). Our experiments thus broaden the base of evidence for the impressive stability of the representation of multiple dynamic objects. Only when frequency of presentation of onset distractors within one trial was dramatically increased - from one in the first three experiments to three to four per second in Experiment 4 - was tracking affected. Even then loss of tracking performance was considerably small.

Alvarez et al. (2005) suggested a two-dimensional spatial memory model (parallel memory access model) as the backbone of the robustness of the tracking system against interruption. We assume that such a memory system as an additional tracking mechanism can offer plausible explanations for our tracking data. On basis of our findings, we assume that a current representation of the state of the relevant objects in this theoretical memory component is constantly available. A tracking system can therefore continue where it left at any given time - even though it has been drawn away from the task unforeseeably. Models of MOT that include more than one process (e.g. Doran & Hoffman, 2010; Fencsik, Klieger, & Horowitz, 2007; Horowitz et al., 2007; Oksama & Hyönä, 2008; Störmer et al., 2013) do explicitly assume an additional routine that manages tracking of multiple items. A plausible explanation for the stability of MOT against interruptions on the basis of such models would be that attention as one process can be captured while the other routine continues keeping spatiotemporal coordinates ready to guide attention back to the designated objects, thus leaving MOT unharmed by sudden disruptions. Although the time span for which the attentional mechanism is absent from the targets is relatively short, a routine that effectively reassigns it to the relevant stimuli is nevertheless necessary. It seems likely that it was the attentional process and not a basic mechanism like the FINSTs that was captured as performance was unaffected in this experiment and we assume such basic processes to constantly provide the necessary dynamic location information. Thus, although online allocation of attention in a dynamic task was effectively influenced by sudden presentation of irrelevant salient distractors, tracking itself was interrupted only when the attentional re-engagement was constantly disturbed, as in Experiment 4.

Our study touches another aspect of attention: The onset distractors in the current experiments were irrelevant, so attending to them would have been orthogonal to the task. It is of foremost importance to monitor “historical continuity” (Pylyshyn, 2004, p. 804) in a tracking paradigm. Therefore, stimuli that onset suddenly beside the relevant target locations were of no importance. Our findings thus also bear some relevance to the discussion of attentional control (see also Burnham, 2007, and Van der Stigchel, Belopolsky, Peters, Wijnen, Meeter, & Theeuwes, 2009, for reviews). Criteria for automatic information processing, as discussed by Yantis and Jonides (1990) and Neo and Chua (2006), are the load-insensitivity criterion and the intentionality criterion. Experiments 1 and 4 clearly established attentional capture in a MOT paradigm. Experiment 2 contrasted with findings of attentional capture in static tasks like visual search, in which manipulation of task load influenced the extent of the attentional capture effect (Cosman & Vecera, 2009; Cosman & Vecera, 2010a; Forster & Lavie, 2008a; Forster & Lavie, 2008b). Although increasing object number from eight to 10 objects affected task difficulty, Experiment 2 replicated the attentional capture effect seen in Experiment 1. Attentional capture in MOT with salient onset distractors thus seems to fulfill the load-insensitivity criterion. In accordance with the study of Neo and Chua (2006), we also found frequency of the onset cue to affect the attentional capture effect (Experiment 3). In contrast to Neo and Chua, the attentional capture effect was found in our experiments in all frequency conditions. Thus, the intentionality criterion may not be violated, although further studies are necessary to thoroughly examine this claim. The attentional capture found in our experiments was "not subject to voluntary control" (Yantis & Jonides, 1990, p.122). Thus, it might be feasible to speak of the effect as an automatic reaction (see also Rauschenberger, 2003, and Theeuwes, 2010, for a discussion of stimulus-driven capture; see Erlikhman et al., 2013, for automatic processes in MOT), although not strongly automatic (Kahneman & Treisman, 1984).

Conclusion

Our findings demonstrate that irrelevant salient onset stimuli are processed during tracking of multiple objects. In the consequence, attention is drawn away from the current focus to these irrelevant stimuli. However, these stimuli are categorized rapidly as irrelevant and attention is quickly disengaged. A continuously updated - probably basic - mechanism provides the necessary location information to guide attention back to the relevant signal. A conclusion our paradigm allows is that besides this basic mechanism there is some kind of an attentional focus on targets in MOT. The absence of this focus leads to an interruption of stimulus processing at the location of the targets. Although in the current form our paradigm does not inform about the exact nature of this attentional focus or the hypothesized basic mechanism involved in tracking, it nevertheless offers a solid prove - although ex negativo - that they are at work at a given moment of time.

Keywords: visual attention, multiple object tracking (MOT), attentional capture, onset distractor

Acknowledgments

The authors thank the two reviewers of this study for their valuable comments and suggestions.

Stimulus images used in the experiments described in this article appear courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, http://www.tarrlab.org/.

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

All data and code used for the analysis for Experiments 1, 2, 3, and 4 are available at PsychArchives: http://dx.doi.org/10.23668/psycharchives.4890 (Dataset) and http://dx.doi.org/10.23668/psycharchives.4889 (R-Script). Example trials are available at: http://dx.doi.org/10.23668/psycharchives.4887.

Commercial relationships: none.
Correspondence: Sebastian Pichlmeier.
Email: sebastian.pichlmeier@ph-karlsruhe.de.
Address: Institute of Psychology, Karlsruhe University of Education, Erzberger Strasse 119, 76133 Karlsruhe, Germany.

Footnotes

[1]All data and code used for the analysis for Experiments 1, 2, 3, and 4 are available at PsychArchives: http://dx.doi.org/10.23668/psycharchives.4890 (Dataset) and http://dx.doi.org/10.23668/psycharchives.4889 (R-Script). Go to http://dx.doi.org/10.23668/psycharchives.4887 for examples of each trial type.
[2]Following Langsrud (2003), the more conservative type II sums of squares were used in this case.

References

Agaoglu, S., Breitmeyer, B., & Ogmen, H. (2018). Effects of Exogenous and Endogenous Attention on Metacontrast Masking. *Vision, 2*(4), 39.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanisms. *Journal of Vision, 7*(13):14, 1–10.

Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception & Performance, 31*(4), 643–667.

Alvarez, G. A., & Scholl, B. J. (2005). How does attention select and track spatially extended objects? New effects of attentional concentration and amplification. *Journal of Experimental Psychology: General, 134*(4), 461–476.

Awh, E., Belopolsky, A., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443.

Bae, G., & Flombaum, J. (2012). Close encounters of the distracting kind: Identifying the cause of visual tracking errors. *Attention, Perception, & Psychophysics, 74*, 703–715.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition, 10*, 949–963.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379–384.

Beck, D., & Lavie, N. (2005). Look here but ignore what you see: effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 592–607.

Bettencourt, K. C., & Somers, D. C. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of Vision, 9*(7):9, 1–11.

Burnham, B. R. (2007). Displaywide visual features associated with a search display's appearance can mediate attentional capture. *Psychonomic Bulletin & Review, 14*, 392–422.

Buschman, T. J., & Kastner, S. (2015). From Behavior to Neural Dynamics: An Integrated Theory of Attention. *Neuron, 88*, 127–144.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*, 1484–1525.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*(7), 349–354.

Cavanagh, P., Holcombe, A. O., & Chou, W. (2008). Mobile computation: spatiotemporal integration of the properties of objects in motion. *Journal of Vision, 8*(12), 1–23.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., . . . De Rosario, H. (2020). *Basic Functions for Power Analysis*. R package version 1.3-0. Retrieved from, https://github.com/heliosdrm/pwr.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98–101.

Cosman, J. D., & Vecera, S. P. (2009). Perceptual load modulates attentional capture by abrupt onsets. *Psychonomic Bulletin & Review, 16*, 404–410.

Cosman, J. D., & Vecera, S. P. (2010a). Attentional capture by motion onsets is modulated by perceptual load. *Attention, Perception, & Psychophysics, 72*, 2095–2105.

Cosman, J. D., & Vecera, S. P. (2010b). Attentional capture under high perceptual load. *Psychonomic Bulletin & Review, 17*, 815–820.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Doran, M. M., & Hoffman, J. E. (2010). The role of visual attention in multiple object tracking: Evidence from ERPs. *Attention, Perception, & Psychophysics, 72*(1), 33–52.

Dux, P., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception, & Psychophysics, 71*(8), 1683–1700.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4*, 345–352.

Erlikhman, G., Keane, B., Mettler, E., Horowitz, T., & Kellman, P. (2013). Automatic feature-based grouping during multiple object tracking. *Journal of Experimental Psychology: Human Perception & Performance, 39*(6), 1625–1637.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Fencsik, D., Klieger, S., & Horowitz, T. (2007). The role of location and motion information in the tracking and recovery of moving objects. *Perception & Psychophysics, 69*, 567–577.

Feria, C. S. (2012). The effects of distractors in multiple object tracking are modulated by the similarity of distractor and target features. *Perception, 41*(3), 287–304.

Flombaum, J. I., Scholl, B. J., & Pylyshyn, Z. W. (2008). Attentional resources in tracking through occlusion: The high-beams effect. *Cognition, 107*(2), 904–931.

Folk, C. L., Remington, R.W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance, 18* (4), 1030–1044.

Forster, S., & Lavie, (2008a). Attentional capture by entirely irrelevant distractors. *Visual Cognition, 16*(2-3), 200–214.

Forster, S., & Lavie, N. (2008b). Failures to Ignore Entirely Irrelevant Distractors: The Role of Load. *Journal of Experimental Psychology: Applied 14*(1), 73–83.

Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science, 21*, 920–925.

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.

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2017). Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception & Psychophysics, 79*, 45–62.

Gaspelin, N., Ruthruff, E., & Lien, M.-C. (2016). The problem of latent attentional capture: Easy visual search conceals capture by task-irrelevant abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 42*(8), 1104–1120.

Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set for displaywide visual features. *Journal of Experimental Psychology: Human Perception & Performance, 24*(3), 699–706.

Horowitz, T. S., Birnkrant, R. S., Fencsik, D. E., Tran, L., & Wolfe, J. M. (2006). How do we track invisible objects? *Psychonomic Bulletin & Review, 13*, 516–523.

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.

Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. (2010). Eye movements across viewpoint changes in multiple object tracking. *Visual Cognition, 18*(9), 1368–1391.

Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition, 20*(7), 848–864.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision, 9*, 1–12.

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. In J. B. Long, & A. D. Baddeley (Eds.) *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman, R. Davies, & J. Beatty (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, nonpredictive function. *Cognitive Psychology, 52*, 346–368.

Kim, M.-S., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of the interference from a salient distractor. *Perception & Psychophysics, 61*(6), 1009–1023.

Lamy, D., Tsal, Y., & Egeth, H. E. (2003). Does a salient distractor capture attention early in processing? *Psychonomic Bulletin & Review, 10*(3), 621–629.

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use type II instead of type III sums of squares. *Statistics and Computing, 13*(2), 163–167.

Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review, 12*(4), 669−674.

Lavie, N., & Fox, E. (2000). The role of perceptual load in negative priming. *Journal of Experimental Psychology: Human Perception and Performance, 26*(3), 1038–1052.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339–354.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual-attention. *Perception & Psychophysics, 56*, 183–197.

Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. R package version 4.4-0. Retrieved from, https://CRAN.R-project.org/package=ez.

Liao, H. I., & Yeh, S. L. (2013). Capturing attention is not that simple: different mechanisms for stimulus-driven and contingent capture. *Attention, Perception & Psychophysics, 75*(8), 1703–1714.

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2020). Progress toward resolving the attentional capture debate. *Visual Cognition*, *29*(1), 1–21.

Makovski, T., & Jiang, Y. (2009). Feature binding in attentive tracking of distinct objects. *Visual Cognition, 17*(1–2), 180–194.

Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2017). Studying visual attention using the multiple object tracking paradigm: A tutorial review. *Attention, Perception, & Psychophysics, 79*(5), 1255–1274.

Neo, G., & Chua, F. K. (2006). Capturing focused attention. *Perception and Psychophysics, 68*(8), 1286–1296.

Ogawa, T., & Komatsu, H. (2004). Target selection in area V4 during a multidimensional visual search task. *Journal of Neuroscience, 24*(28), 6371–6382.

Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology, 56*(4), 237–283.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 2–25.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*(2), 160–174.

Potter, M. C., Staub, A., & O'Connor, D. H. (2002). The time course of competition for attention: Attention is initially labile. *Journal of Experimental Psychology: Human Perception & Performance, 28*(5), 1149–1162.

Pylyshyn, W. Z. (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. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition, 14*(2), 175–198.

Pylyshyn, Z., Haladjian, H., King, C., & Reilly, J. (2008). Selective nontarget inhibition in multiple object tracking (MOT). *Visual Cognition, 16*(8), 1011–1021.

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

Rauschenberger, R. (2003). Attentional capture by auto- and allo-cues. *Psychonomic Bulletin & Review, 10*(4), 814–842.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, https://www.R-project.org/.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217–236.

Ruz, M., & Lupiañez, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control. *Psicológica, 23*, 283–309.

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. 49–77). Cambridge, MA: MIT Press.

Scimeca, J. M., & Franconeri, S. L. (2015). Selecting and tracking multiple objects. *Wiley Interdisciplinary Reviews: Cognitive Science 6*, 109–118.

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. *Journal of Neuroscience, 33*(12), 5346–5351.

St. Clair, R., Huff, M., & Seiffert, A. E. (2010). Conflicting motion information impairs multiple object tracking. *Journal of Vision, 10*(4):18, 1–13.

Theeuwes, J. (2010) Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99.

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition, 1*(1), 29.

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In S. Monsell, & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 105–124). Cambridge, MA: MIT Press.

Theeuwes, J., De Vries, G.-J., & Godijn, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics, 65*(5), 735–746.

Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin & Review, 11*, 697–702.

Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple object tracking. *Cognition, 108*, 1–25.

Torchiano, M. (2020). Efficient Effect Size Computation. R package version 0.8.1, https://github.com/mtorchiano/effsize/.

Tran, A., & Hoffman, J. (2016). Visual attention is required for multiple object tracking. *Journal of Experimental Psychology: Human Perception & Performance, 42*(12), 2103–2114.

Van der Stigchel, S., Belopolsky, A. V., Peters, J. C., Wijnen, J. G., Meeter, M., & Theeuwes, J. (2009). The limits of top-down control of visual attention. *Acta Psychologica, 132*(3), 201–212.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.

Yantis, S. (2000). Goal-Directed and Stimulus-Driven Determinants of Attentional Control. In S. Monsell, & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 73–103). Cambridge, MA: MIT Press.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601–621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance, 16*(1), 121–134.

Yeshurun, Y., Rashal, E., & Tkacz-Domb, S. (2015). Temporal crowding and its interplay with spatial crowding. *Journal of Vision, 15*(3):11, 1–16.

Zhao, L., Gao, Q., Ye, Y, Zhou, J., Shui, R., & Shen, M. (2014). The role of spatial configuration in multiple identity tracking. *PLoS One 9*(4), e93835.