Atten Percept Psychophys (2016) 78:2090–2109
DOI 10.3758/s13414-016-1129-0

All eyes on relevance: strategic allocation of attention as a result of feature-based task demands in multiple object tracking

Alisa Brockhoff[1] · **Markus Huff**[1]

Abstract Multiple object tracking (MOT) plays a fundamental role in processing and interpreting dynamic environments. Regarding the type of information utilized by the observer, recent studies reported evidence for the use of object features in an automatic, low- level manner. By introducing a novel paradigm that allowed us to combine tracking with a noninterfering top-down task, we tested whether a voluntary component can regulate the deployment of attention to task-relevant features in a selective manner. In four experiments we found conclusive evidence for a task-driven selection mechanism that guides attention during tracking: The observers were able to ignore or prioritize distinct objects. They marked the distinct (cued) object (target/distractor) more or less often than other objects of the same type (targets /distractors)—but only when they had received an identification task that required them to actively process object features (cues) during tracking. These effects are discussed with regard to existing theoretical approaches to attentive tracking, gaze-cue usability as well as *attentional readiness*, a term that originally stems from research on attention capture and visual search. Our findings indicate that existing theories of MOT need to be adjusted to allow for flexible top-down, voluntary processing during tracking.

Keywords Multiple object tracking · Multiple identity tracking · Attentional control · Attentional resolution · Attention allocation · Goal-directed attention

Introduction

In dynamic real-life environments, keeping track of relevant objects can be a rather challenging task. For example, in crowded inner-city pedestrian areas, multiple salient cues (ads, blinking lights, etc.) attract visual attention that might distract from keeping track of family members (e.g., children or the dog): stimulus-driven processes interact with goal-directed tasks during attentive tracking. In the present manuscript, we explore how these processes affect the distribution of visual attention in dynamic scenes and contribute to the human ability of multiple object tracking (MOT).

Attention and focus

It is a well-known fact that attention is influenced actively or passively (James, 1890). We can actively exert control over the allocation of attention (top-down; e.g., Yarbus, 1967) or the deployment of attention happens passively as a result of an event in the environment (bottom-up; e.g., Pratt, Radulescu, Guo, & Abrams, 2010). The question of how narrowly attention can be focused has been characterized as a question of attentional resolution. Based on different metaphors that described visual attention (including attention as a spotlight: Posner, 1980; as a nonlinear filter: Cutzu & Tsotsos, 2003; as a zoom lens: Eriksen & James, 1986), researchers wondered about the limitations of spatial extent of focused attention and suggested that the area of selection has a facilitatory-center-inhibitory-surround profile (e.g., Cutzu & Tsotsos, 2003). A detailed discussion of all approaches and findings in the area of attention selection, facilitation, and inhibition would go beyond the scope of this work. However, a finding that is important for the studies we report here is that the shape of the focus varies depending on task and characteristics of stimuli (e.g., Eriksen & James, 1986; Laberge & Brown,

✉ Alisa Brockhoff
alisa.brockhoff@uni-tuebingen.de

[1] Department of Psychology, University of Tübingen, Tübingen, Germany

1986). Eriksen and James (1986) built upon existing research methods in which participants were asked to discriminate between the letters *S* and *C* as quickly as possible. In their static display they manipulated the number of cued positions and the discriminative difficulty with incompatible noise letters, and measured reaction times. Their gathered data indicated that participants were able to change the size of the attentional focus and drew upon additional attentional resources when the number of cued positions increased. The authors interpreted their findings as support for the zoom lens model.

Similar theoretical approaches were postulated later. Morgan, Ward, and Castet (1998), and Morgan and Watt (1997) suggested that cues do not elicit a switch to a higher spatial frequency passband, but act upon the scale of spatial frequency analysis. Thus, focusing attention (e.g., with a cue) allows processing with a smaller spatial range, resulting in a decreased influence of distractors and increased target acuity.

Attention and cognition

Observed effects of selective and focused attention in several cognitive tasks also reflect a combination of different mechanisms such as noise reduction, signal enhancement, and decisional as well as intentional factors (Folk, Remington, & Johnston, 1992; Lu & Dosher, 1998; Morgan et al., 1998; Yeshurun & Carrasco, 1998, 1999). The compelling demonstrations of both early and late selection imply that attention is located flexibly (see also Vogel, Woodman, & Luck, 2005).

Several authors have argued further that passive/exogenous signals can only capture attention when the observer has an optimal attentional control set (ACS; Folk et al., 1992). Yantis and Egeth (1994) reported that response times to singleton targets are highly sensitive to the relevance of the singleton (see also Folk & Annett, 1994; Hillstrom and Yantis, 1994; Jonides & Yantis, 1988). That is, it highly depends on the observer's "attentional readiness" (top-down or goal-directed control, as described in Egeth & Yantis, 1997) or its equivalent, the ACS (as used as the generic concept for attentional readiness throughout the rest of this paper; e.g., Folk et al., 1992; Folk, Remington, & Wright, 1994) to find an *L* in an array of *T*s (Joseph & Optican, 1996). Others also recently argued that, although exogenous cues might work to orient attention in space, the *strength* of the effect may be endogenously modulated by an ACS that the observer adopted in line with a given goal or task (Lupiáñez, Milliken, Solano, Weaver, & Tipper, 2001).

Attention and spatial resolution

The widely accepted view is that attention is used to control the allocation of limited perceptual processing resources in various ways. The described studies present evidence that attention can increase or decrease its resolution as a result of spatial occurrences (e.g., crowding). Furthermore, attention changes the range of spatial analysis (e.g., cueing). And finally, attention itself is modified by task and intention of the observer (e.g., attention capture and control settings). Since the first statement concerning *spatial* resolution has already been explored in dynamic environments (e.g., Franconeri, Jonathan, & Scimeca, 2010; Intriligator & Cavanagh, 2001), the current paper is concerned with the effects of a flexible attention allocation and cue usability that result from *feature*-based task demands in a multiple object tracking paradigm (MOT; Pylyshyn & Storm, 1988).

Interestingly, while the common understanding of resolution (e.g., of a TV or a picture) rather describes how many details the observer can see, attentional resolution during tracking is usually associated with locations and the selection and separation of objects (see Intriligator & Cavanagh, 2001), and not with the perception of object features. From a technical point of view, resolution is measured as how closely lines can be resolved in an image, and the clarity of an image depends in fact on its *spatial* resolution (and not as, often erroneously believed, on the number of pixels per inch, ppi). The current paper therefore adopts, and by that extends, the term attentional resolution to describe how a task changes attentional resolution to represent object features with more clarity during MOT. Following Luck, Hillyard, Mouloua, and Hawkins (1996)—who proposed that there are multiple selection mechanisms that operate at different processing levels to control different types of interference and attentional overload (see Cave & Bichot, 1999, for a review)—we expect the pattern of attentional resolution to change, depending on the mechanism triggered by the demands of the task. That is, we propose that a featural task demand will trigger and activate a feature level of attention that influences the allocation and resolution of attention for object features in a dynamic tracking environment.

Attention and MOT

A typical MOT experiment presents the subject with a number of identical objects of which some are briefly flashed to indicate that they are the targets to be tracked. Then, the objects become indistinguishable from each other and move. When the motion stops, the observer is asked to click on the objects that were marked as targets before. The MOT task has been originally invented to show that observers track objects in parallel following a preattentive mechanism (Pylyshyn, 1999). As described by Pylyshyn (2007), visual attention uses indexes that stick to the moving targets, and these indexes are attracted to moving objects in a bottom-up manner.

However, the described indexes serve the purpose of providing a structure for guiding focused attention needed in specific situation, for example, in situations of object crowding in which the observer has to prevent confusions. That is, the basic MOT mechanism is a preattentive "cognitively impenetrable" mechanism (Pylyshyn, 1999) but may pave the way for top-down influences. Still, the role of visual attention in tracking has been approached in several different theoretical ways, and no agreement has been reached on the issue. Cavanagh and Alvarez (2005) proposed that multifocal attention covers the objects simultaneously and independently. Alternatively, Alvarez and Franconeri (2007) suggested that a limited attentional resource is allocated flexibly toward objects, increasing local attentional resolution and enhancing tracking performance. Iordanescu, Grabowecky, and Suzuki (2009) supported the flexible idea with their study on the dynamic adjustment of the spatial distribution of attention in an MOT environment based on spatial demands (e.g., crowding). They blanked out the MOT task and asked participants to localize the position of targets on the blank screen. Their results, namely that the accuracy of correctly marked targets increased when the distance of the nearest distractor decreased, indicated that the attentional allocation to individually tracked target objects changes dynamically. That is, close distractors lead to an increased attentional resolution (e.g., to prevent confusions), or to a modulation of the local attentional resolution, possibly by a top-down component, as you will. Increasing or decreasing the attentional resolution based on current demands has only been investigated for spatiotemporal occurrences, like crowding or overlapping objects, but has never been investigated for goal-related processing of feature information.

Location, features, and MOT

Concerning top-down effects for object features during tracking, Feria (2012) was the first to study the effects of distinct object features on tracking performance. The utilized distractors were either identical or featurally distinct (in one or two dimensions) from the targets; the number of distractors per trial was varied. Her findings indicated that the effect distractors have on tracking is top-down in nature: differently colored or shaped, or motionless distractors impaired tracking less than target-identical distractors—still, this was only the case when tracking load was low. With her study she generalized previous findings from visual search to MOT: The effect of a distractor object is dependent on sharing the features of the target, indicating that the role of distractors as well as distractor features may have been underestimated in its influence on tracking performance.

Still, most theories on MOT have a strong focus on spatiotemporal information as the only source used (see, e.g., Cavanagh & Alvaraz, 2005; Oksama & Hyönä, 2008; Pylyshyn, 1989). Nonetheless, a considerable amount of tracking studies has found effects that can be attributed to information access to features during tracking (e.g., Cohen, Pinto, Howe, & Horowitz, 2011; Drew, Horowitz, & Vogel, 2013; Makovski & Jiang, 2009a, b; Oksama & Hyönä, 2008). While typically irrelevant for tracking (Makovski & Jiang, 2009a, b; Pylyshyn, 2004), features are used when spatiotemporal information is(made) unreliable (Bae & Flombaum,2012; Papenmeier, Meyerhoff, Jahn, & Huff, 2013). This indicates again, that attention to different sources of information (here: object's feature or location) is somewhat flexibly.

MOT, features, and attentional resolution

With the present research, we explore the role of top-down modulation of attention allocation and feature processing in MOT. We hypothesize that an increased or decreased attentional resolution can also happen as a response to a featural demand, shifting attention toward or away from a feature singleton within an otherwise homogenous crowd of objects. Following the idea of an ACS, we hypothesize that based on its relevance for the tracking task, target and distractor singletons will lead to changes in the allocation of attention; that is, an intentional focus on feature singletons will lead to a strategic use of helpful target singletons and an inhibition of tracking-harming distractor singletons. While being aware however that our study does not explicitly test attention capture, we are convinced that a task-related manipulation of the relevance for features in a dynamic environment bares a practical share to the induction of an ACS used in attention capture settings. To understand the factors that control attentional resolution and possible effects of accessing object features, it is important to be able to systematically vary the locus of selection within a single paradigm. We present here a new method of manipulating the attentional resolution toward a single object feature in an MOT task. By using dynamic gaze cues (cartoon eyes with moving pupils), that cued either a single target or a single distractor among the objects by looking at them, we integrated minimal featural gaze cues within an MOT task (see Fig. 1).

The main objective was to enhance the attentional resolution to allow for feature processing by influencing the allocation of attention. In three of the four experiments presented here, seven dynamic cartoon eyes cued one neutral pair of eyes by gazing at it. Such a "dual function" of objects—that is, the observer can process the objects by following the spatial cues, or by processing featural distinctiveness, or both—was inevitable in a dynamic cueing design. The combination

Fig. 1 Example of the stimuli. Left: Eyes with neutral gaze. Middle: Eyes cueing one of the objects. Right: Eyes with color cue (used in Experiment 4).

of gaze cues and MOT had an exploratory character with two possible outcomes: Either participants use the cue itself reflexively, or they only focus on the different featural aspects of the objects, or both.

We implemented two gazing behaviors in order to control for the possible difference between feature encoding and cue processing, and to avoid difficulties in the later interpretation of the results. The gaze cue we presented was either constant ("stalking") or intermittent ("flirty"). The stalking condition presented stimuli with constant orientation toward the cued object and thus gave the participants 8 full seconds to encode object features. The flirty condition however presented a short gaze twice during a trial and we assumed that feature encoding would be very difficult, that is, we hoped to observe reflexive gaze-following effects.

Attention and gaze

The rational behind introducing the gazing behavior of the cartoon eyes was that gaze perception can direct attention reflexively, performed by an innate module (Eye Direction Detector; Baron-Cohen, Campbell, Karmiloff-Smith, Grant & Walker, 1995) and also known as joint attention, mainly studied in infants (e.g., Farroni, Massaccesi, Pividori, & Johnson, 2004; Scaife & Bruner, 1975). Findings indicated that infants followed the eyes of the speaker and turned their attention to the looked-at object, a behavior that promotes the acquisition of language (e.g., Baldwin, 1992, 1993). The spatial cueing paradigm and its variations (Posner, 1980; Posner & Cohen, 1984; Posner, Nissen, & Ogden, 1978) showed that gaze keeps playing a strong role for the orientation of attention throughout the life span. Friesen and Kingstone (1998) studied the effects of gaze cues on the orienting of attention in adults. Presenting faces in the center of the screen whose eyes either gazed right or left resulted in facilitated reaction times when the target appeared in the gaze-cued area.

Gaze and attentional control

However, Driver et al. (1999) informed participants in one of their experiments that the target was four times more likely to appear at the noncued side. In contrast to the assumption of an automatic, reflexive effect of gaze, they found that participants eventually shifted their attention voluntarily in the opposite direction of the cue. Furthermore, a recent study by Macdonald and Tatler (2013) provided insights into how gaze cues are actually used in real-life scenarios. Participants had to build a given structure with colored blocks, receiving either ambiguous or unambiguous instructions. They found that participants only used the gaze cues of an attending experimenter when instructions were ambiguous, that is, when the gaze cue provided information that was helpful to solve the task. This suggest a strong influence of task demands on the effects of gaze cues.

Experimental overview

In three of the four studies presented here we exploit the property that spatial cues focus attention on an area in visual space and that the allocation of attention has a large effect on feature detection and encoding (Treisman & Sato, 1990; Theeuwes, Kramer, & Atchley, 1999; Treisman & Gormican, 1988). However, this has only been tested with static displays and visual search so far.

In Experiment 1, participants tracked one round without further information. Before the second round of tracking, they received the task to identify the objects' behavior (cueing constantly or intermittently) and, additionally and more importantly, to identify the cued object (target or distractor singleton). Instead of measuring proportion correct, we measured the selection preference (proportion marked) for the different object types displayed: targets, distractors, and the singleton (cued distractor or cued target), to provide evidence that facilitated feature processing based on task demands influences the allocation of attention flexibly among the objects. This approach is then used in Experiments 2 and 3 to demonstrate that the effect is truly task-based and not due to inevitable learning. Using the same methodological design with different stimuli, we show in Experiment 4 that the use of featural gaze-cues is basically identical to color-cues, indicating that cue usability in MOT relies upon the same flexible and intentional attentional enhancement for object features.

General method

Overview

In the experiments reported here, we used the MOT task developed by Pylshyn and Storm (1988) in combination with an identification task. In three of the four experiments, there were two rounds of tracking. Round 1 was with no task or further information. In Round 2 participants received information on the objects' behavior and the task to identify it after each tracking trial. In Experiment 2, we added a third round in which participants still knew about the eyes' behavior but were explicitly informed that there will be no identification task.

Stimuli

Two types of stimuli were used. In the first three experiments, we used cartoon eyes that gazed at one of the target, or one of the distractor objects (see Fig. 1). In Experiment 4, we used the same eyes but instead of a gaze all objects except the single target or the single distractor were colored grey (see Fig. 1). The cue was either displayed constantly (*stalking*) or intermittently (*flirty*); 500 ms before the motion stopped, all objects went back to their neutral initial position.

Phase 5 (neutral eyes) was displayed until the end of the trial in *flirty* trials. In *stalking* trials, the eyes switched to the neutral position 500 ms before they stopped moving. Thus, in both conditions it was not sufficient to pay attention to the cues at the end of the trial only. Correct identification of the objects' behavior and the cued object required sustained attention throughout the entire trial.

Procedure

At the start of each session, participants received the Reading the Mind in the Eyes test (Baron-Cohen et al., 2001; Bölte, 2005). This was followed by instructions for one round of tracking (64 trials). Except for Experiment 3 (tracking without an additional identification task), participants then received specified instructions on the eyes' behavior, and a multiple-choice question appeared on the screen after each trial. In Experiment 2, participants tracked an additional round without a task. The tracking procedure was a standard procedure, which is depicted in Fig. 2, which also depicts a *flirty* trial.

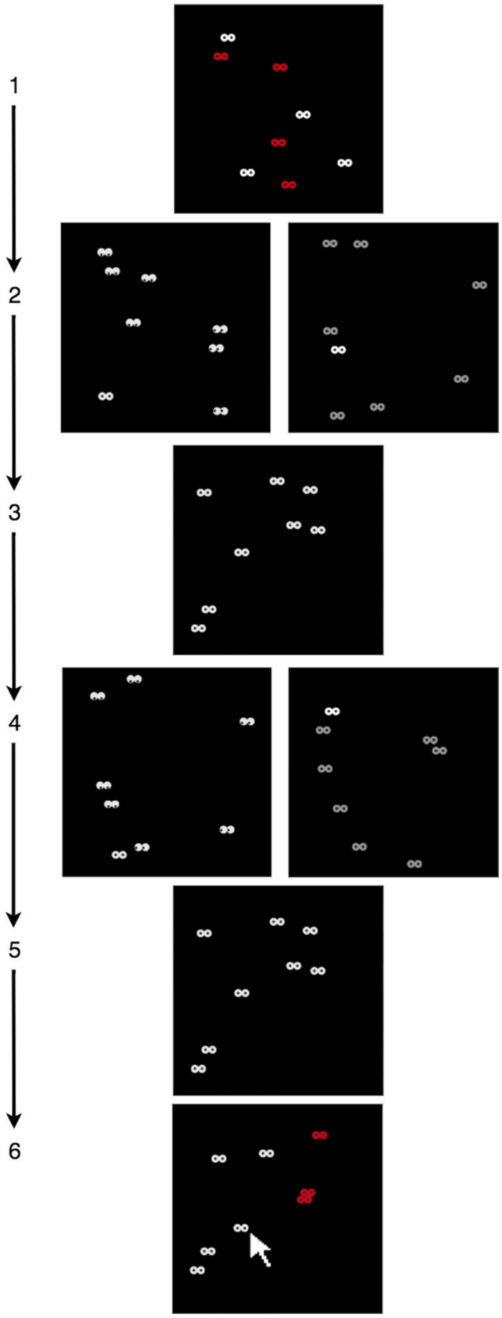

Fig 2 Example of a *flirty* trial. Left: gazing eyes. Right: Shaded eyes. Numbers indicate the sequential differences displayed. Note that the flirt moments (2 and 4) appeared randomly at 2 of 3 predefined moments during the trial. In *stalking* trials, the cueing phase was constant and lasted through Sequences 2 to 4.

Exclusion criteria and data analysis

In each of the experiments described in this work, we tested participants on their eye-reading skills with the Read the Mind in the Eyes test by Baron-Cohen et al. (2001). It seems counterintuitive that a test that measures the ability to correctly judge an emotion from gaze is appropriate to use as a baseline for the sensitivity to gaze cues that are directional in nature. However, recent findings suggest that gaze direction and facial expression are not independently processed; in fact, gaze direction is considered as an important cue in the perceptual processing of facial displays of emotions (e.g., Adams & Kleck, 2003). Further, the Read the Mind in the Eyes test is frequently used (as part of a test battery) in clinical research and considered to be equally reliable for participants in treatment as well as for healthy participants in control groups. Reflexive orienting to gaze cues can be reduced or absent in people who suffer from psychiatric disorders (e.g., Asperger's or autism; see Nation & Penny, 2008, for a review). Thus, with relatively simple means, we attempted to ensure a rough comparability of sensitivity to gaze cues among participants. Additionally, data of trials in which the participant did not correctly identify the condition (see Appendix A for chi-square tests of the differences in proportion correct), that is, failed to process the task, were excluded from the analysis. Such an exclusion of data was justifiable from a theoretical point of view: we needed to ensure that attention is actually controlled by top-down processes (task). Considering the goal of the study, namely to explore the influence of "active" task-related attention, we believe that an inclusion of trials in which it is impossible to tell what actually happened during tracking (e.g., the participant only tracked the objects without trying to identify the condition and guessed, or the participant accidentally tracked a distractor and came to wrong conclusions when asked which object has been cued) bares the risk to lead to false conclusions based on unknown and uncontrolled factors in the analysis.

The planned contrasts were kept within one gazing behavior (*flirty* and *stalking*), that is, we only compared the proportion of the selection of noncued objects versus cued objects in *flirty* or *stalking* trials. In other words, we tested *a subset of possible main effects* for the type of object (cued vs. noncued), and none of our comparisons involved both the *flirty* and the *stalking* trials, nor *Round 1* and *Round 2* trials. In the fifth contrast, though we analyzed the difference of the differences of each behavior. However, due to differences in proportions correctly identified we did not give this calculation too much weight in the result section. Based on our specific a priori expectations that noncued objects would be marked less often than cued objects when the participants' attention was influenced by task-demands, an omnibus *F* test like a repeated-measures ANOVA and subsequent pairwise comparisons would have resulted in an inflated Type I error and were thus not the most appropriate analysis (see Maxwell & Delaney, 2004; Ruxton & Beauchamp, 2008). We chose planned contrasts that derived from specific hypotheses (see Table 1). We further applied Bonferroni corrections to reduce the risk of type I errors and to compensate for the fact that our set of chosen contrasts was not orthogonal.

Experiment 1

Method

Participants

Forty students (28 female, 12 male; mean age $M = 22.63$ years, $SD = 3.32$), of the University of Tübingen participated in return for course credit or monetary compensation. All subjects had reportedly normal or corrected-to-normal sight. Informed consent was obtained after the experimental procedures were explained to the subjects.

Stimuli and design

PsychoPy (Peirce, 2009) was used to present stimuli on a 15.4-in. notebook. The experiment was divided into three parts. The participants started with a revised version of the Reading the Mind in the Eyes test (Baron-Cohen et al., 2001; Bölte, 2005) that took them approximately 10 minutes to complete. Stimuli consisted of 37 photographs of human eye-regions, the first one being a practice trial. These photographs were taken from the revised version of the Reading the Mind in the Eyes test (Baron-Cohen et al., 2001) translated into German by Bölte (2005). Four different adjectives were given. The participant indicated the adjective that described the mental and emotional state of the person displayed best.

Next, participants were given instructions for a MOT task. Each trial started with the appearance of rectangular outlined space against a black background. After 0.5 seconds, the stimuli appeared. After another 0.5 seconds, four of the stimuli were marked as targets by flashing red for four times. Each flash and each pause in between lasted for 0.2 seconds, while the last flash lasted for 0.4 seconds. The stimuli were eight artificial, cartoon eyes with a white sclera and black pupils and had a diameter of 2.6° × 1.3° (see Fig. 1). They moved at a constant speed of 8°. The duration of a trial was 8 s, and participants were instructed to select the target items by marking them after each trial. All objects were able to overlap during each trial and they bounced off of the box in a way that was physically correct. Gaze direction was manipulated by moving the black pupils as follows: The eyes either constantly stared ("stalking") at a specified object or looked at it shortly ("flirted") for twice during each trial.

The two "flirt moments" were randomly chosen out of three possible time points (in order to avoid predictability) and lasted about 1.5 s each (that included movement of the pupil toward the object, moment of glance, and the aversion of the eyes). The specified object either belonged to the defined set of targets or was one of the distractors. Each of the four conditions (i.e., the four ways in which the eyes could behave) occurred twice in each of 16 blocks, with two blocks of practice (which were not further analyzed). Participants tracked two rounds of 64 trials, so 128 trials in total. Round 1 was with no task or further information. In Round 2, participants received information on the objects' behavior and the task to identify it after each tracking trial. The first round of tracking took approximately 20 minutes to complete; the second round took 25–30 minutes. Before the second round of tracking, participants were actively engaged in attending to the objects' behavior during tracking (as in Huff, Papenmeier, & Zacks,

Table 1 Research hypotheses for target and distractor trials (Experiment 1, applicable to 2 and 4 as well)

Type of Influence		Selection Preferences	Contrasts
Round 1, Stalking Round 1, Flirty	H_1 H_2	Cued objects will not be marked more often than non-cued objects.	S: CT vs NCT S: CD vs NCD F: CT vs NCT F: CD vs NCD
Round 2, Stalking Round 2, Flirty	H_3 H_4	If participants received information and task, they will be attentionally ready to process features. Cued objects will be marked more often than non-cued objects when they are relevant for the identification task.	S: CT vs NCT S: CD vs NCD F: CT vs NCT F: CD vs NCD
Round 2: Flirty vs Stalking	H_5	We explore the effects of two cue intensities. Differences between the two conditions may be explained with the observers' reaction to the cue (Flirty) or with relying on the distinct object feature that is constantly available (Stalking).	S: CT - NCT vs F: CT – NCT S: CD - NCD vs F: CD - NCD

Notes. S = stalking, F = flirty, CT = cued target, NCT = non-cued targets, CD = cued distractor, NCD = non-cued distractors

2012). The research assistant handed out a written description of the behavior and the resultant answer options (translated):

In the third part of the experiment, you will track pairs of cartoon eyes again. This time you not only have to track the predefined targets, but also have to pay attention to how the eyes behave. There are four options that will be given as a multiple-choice after each tracking trial:

1) The eyes **stared** constantly at a **target** object.
2) They eyes **stared** constantly at a **distractor** object.
3) The eyes **looked twice** quickly at a **target** object.
4) The eyes **looked twice** very quickly at a **distractor** object.

The research assistant stayed in the room to answer individual questions of the participants in case necessary and left before she or he started the second tracking round. The participants were further informed that they would have to rate how confident they were about their answers (5: *maximally confident*–1: *not at all*). The tracking task with questions took about 25 minutes.

Results

With an average of 25.07 correctly identified gaze expressions in the Reading the Mind in the Eyes test, subjects in the present study scored only slightly lower than the normal control group suggested by Baron-Cohen et al. (2001; 26.2–30.9 of 36 gaze expressions in total). No participant was excluded from the analysis. Overall, participants tracked 2.87 of 4 objects correctly. There was no significant effect for *round;* that is, the participant did not track less or more objects in the second round with the identification task than in the first round without an additional task, $t(39) = 0.53$, $p = .598$, $d = 0.08$.

The main analysis focused on the mean proportions of *object selection preferences* (object marked) on the two levels of *object status* and the two levels of *cueing behavior*. We operationalized the ACS as trials in which participants identified the eyes' behavior correctly, resulting in the inclusion of 64 % of the total number of trials of Round 2. The planned contrasts were based on a linear mixed-effects model, fit via maximum likelihood, with *participants* as the random effect and the variables *condition* (Round 1, Round 2), *object status* (Cued, Non-Cued) and *cueing behavior* (Flirty, Stalking). A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .01 per test (.05/5). We report Cohen's d as effect size measures. The planned contrasts for target trials were not significant in Round 1, neither in the *Stalking* condition, $|z| = 0.697$, $p = .928$, $d = 0.01$, nor in the *Flirty* condition, $|z| = 0.22$, $p = .999$, $d = 0.01$. In Round 2, both contrasts reached significance, *Stalking* $|z| = 5.40$, $p < .001$, $d = 0.99$; *Flirty* $|z| = 6.14$, $p < .001$, $d = 1.37$. The distractor trials showed a similar picture for Round 1, but showed no significant differences between the selection of cued distractor and non-cued distractors. See Appendix B for complete tables. Refer to Fig. 3 for a graphical display of the analyzed target distractor trial data.

In the first round, the participants did not process object features, i.e. they did not show a selection preference for the cued target. In contrast, in the second round in which the cued single target was of relevance for both tasks (tracking and answering the *cueing behavior* question), they selected it significantly more often compared to the non-cued targets. The same analysis was done for distractor trials. Here, we would expect that the cued single object (one of the distractors) is either ignored strategically (i.e. marked less often than the non-cued distractors) or not processed differently at all, since both, the tracking task *and* the question, could be solved without guiding attention away from the four target objects.

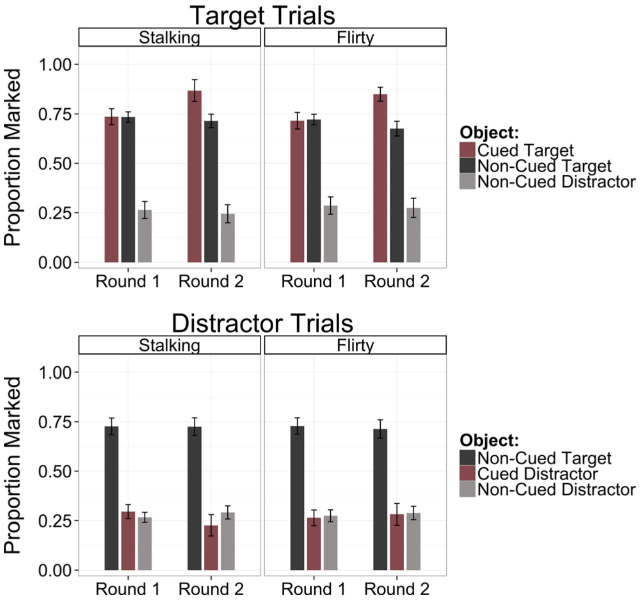

Fig. 3 Proportion marked of the different object types displayed in both rounds for *cued-target* and *cued-distractor* trials depending on *cueing behavior.* Error bars indicate 95 % within-subjects confidence intervals.

Statistically, there is no significant difference between the compared means.

Discussion

The first experiment introduced an ACS *manipulation* before the second round of tracking. Information and task showed strong effects on the selection preference for the cued target but not for the cued distractor. Whether the objects were flirted with or constantly stared at did not influence the strength of the selection preference for cued-target trials. It is important to note, that the results of Round 2 cannot be explained by strategically paying attention only to the very end of each trial. First, 500 ms before the end of a trial all objects were indistinguishable (i.e. the pupils turned into the neutral position). Second, even if there might have been some strategic processing in the *Stalking* condition, the randomly chosen cueing intervals in the *Flirty* condition presumably impeded such a strategy. As the results of the *Flirty* and *Stalking* conditions in Round 2 were comparable, we propose that the ACS did not just trigger a strategy that selects the relevant information when asked for but rather changes the distribution of spatially distributed visual attention during the whole trial. For the cued-distractor trials, we only observed a close to significant inhibition effect for the cued distractor: it was marked less often than the non-cued distractors.

The research assistant stayed in the room to answer questions after the participant had received the instructions for Round 2 and reported that the majority of participants had questions about the explanations on distractor trials. Even though participants scored well over the multiple-choice guessing rate of 25 % (about 65 % of trials correctly identified), we wondered whether the influence of the ACS would change with revised instructions. Considering the indistinct effect for cued-distractor trials and taking the concerns about the instructions into account, we conducted a second experiment in which we handed out an illustrated (as opposed to the former written version) instruction to the participants. Furthermore, a third round of tracking was added in which no identification of the *cueing behavior* was required. That is, in Round 3 the participants still knew about the behavior of the different object types. If results of Round 3 resemble those of Round 2, we could derive that knowledge (without the task) is capable to amend the distribution of attention. But if results of Round 3 rather resemble those of Round 1, the difference in processing of cued objects owes its existence to the task demands that decreases relevance for cued distractors, and heightens relevance for cued target objects. Experiment 2 was thus not only conducted to replicate results of Experiment 1 with modified instructions, but also to consider and extract the different roles task and information play in eliciting the previously observed effect.

Experiment 2

Method

Participants

Thirty-two new students of the University of Tübingen participated in return for course credit (24 female and 8 male, aged $M = 22.42$ years, $SD = 3.64$). All subjects had reportedly normal or corrected-to-normal sight. Informed consent was obtained after the experimental procedures were explained to the subjects.

Stimuli and design

The design of the experiment was similar to Experiment 1, except that we added a third tracking round. Participants tracked 192 (3 × 64) trials in total. The first and the second round were identical to the former experiments: the participants tracked the first round *without* a task, and the second round *with* a task. Before the third round started, the participants were informed that the eyes will move and gaze in the exact same way, but that they will not be asked to identify the eyes' behavior at any point during Round 3. The participants also did the "Read the Eyes in the Mind Test". We excluded no participant based on the test performance. The stimuli were not changed. However, we revised the instructions that were now illustrated with a series of screenshots and detailed descriptions of the real stimuli, one for cued-target and one for cued-distractor trials.

Results

With an average of 25.35 correctly identified gaze expressions in the Reading the Mind in the Eyes test, subjects in the present study scored only slightly lower than the normal control-group suggested by Baron-Cohen et al. (2001; 26.2 - 30.9 of 36 gaze expressions in total). No participant was excluded from the analysis. There was no significant effect for *round* that is the participant did not track less objects in the second round with the identification task than in the first round without an additional task, $t(31) = 0.55$, $p = .58.$, $d = 0.09$. Neither differed the second from the third round, $t(31) = 1.05$, $p = .30.$, $d = 0.18$. Based on participants' correct answers, we included 66.2 % of the total trials of Round 2. The planned contrasts were based on a linear mixed-effects model, fit via maximum likelihood, with *participants* as the random effect and the variables *condition* (Round 1, Round 2, Round 3), *object status* (Cued, Non-Cued) and *cueing behavior* (Stalking, Flirty). A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .006 per test (.05/8). The planned contrasts for target trials were not significant in Round 1, neither in the *Stalking* condition, $|z| = 0.77$, $p = .97$, $d = 0.18$, nor in the *Flirty* condition, $|z|= 0.44$, $p = .99$, $d = 0.12$. Also in Round 3 there were no significant effects, neither in the *Stalking* condition, $|z| = 0.847$, $p = .96$, $d = 0.20$, nor in the *Flirty* condition, $|z| = 1.10$, $p = .87$, $d = 0.26$. In Round 2, both target contrasts were significant, *Stalking* $|z| = 5.08$, $p < .001$, $d = 1.21$; *Flirty* $|z| = 3.87$, $p < .001$, $d = 0.75$. The distractor trials showed a similar picture for Round 1 and Round 3 (no significant effects). However, in Round 2 the difference in the selection of the cued distractor and the non-cued distractors was significantly different from zero in the *Stalking* condition, $|z| = 3.81$, $p < .001$, $d = 0.95$. See tables in Appendix C. Refer to Fig. 4 for a graphical display of the analyzed target and distractor trial data.

The results of cued-*target* trials replicate our findings of Experiment 1. The interaction contrasts of non-cued and cued targets are only significant in Round 2, but neither in Round 1, nor in Round 3 of the tracking task. The results of the cued *distractor* trials were surprising. The revised instructions led to the hypothesized effect of strategic distractor inhibition - but again only for trials in which eyes cued the distractor in a *stalking* manner. The effect was not observed for *flirty* cues. The difference in proportion marked of cued and non-cued objects differed significantly between the two *cueing behaviors*: the neutral eyes were marked more often when stared at constantly compared to looked at only twice.

Discussion

The second experiment replicated the design of Experiment 1 and added a third round of tracking. The findings of the preceding results are reflected in the current data. Under the ACS manipulation, the single cued target-objects were marked more often than the other targets when the participant was attentionally ready. This leads us to assume that the demands of the identification task play a stronger role here than knowledge about features alone.

In Experiment 1, we found a slight hint of an inhibition of the single object when it was a distractor and constantly cued (*Stalking* condition), which is why we introduced revised instructions in Experiment 2 to ensure that participants understood the concept of distractor objects completely. And in fact, the effects found for the cued-distractor objects in Experiment 1 reached statistical significance in the current data. Yet be aware that the distractor object was marked less often (than the three equally available other distractors) *only* in the *Stalking* but not in the *Flirty* condition. While Experiment 2 showed that the effects of the ACS on object selection is only observable when it is activated but is disregarded when the participant is no longer asked to focus on features, it may be possible that the effects observed in Experiment 1 have a learning component. That is, experience in tracking may result in better integration of object features and possible learning for helpful cues without further information. Experiment 3 replicated the design of Experiment 1 but participants were not given any information or an identification task.

Experiment 3

Method

Participants

Twenty students (12 female, 8 male; mean age $M = 21.75$ years, $SD = 3.18$), of the University of Tübingen participated in return for course credit or monetary compensation. All subjects had reportedly normal or corrected-to-normal sight. Informed consent was obtained after the experimental procedures were explained to the subjects.

Stimuli and design

Stimuli and design was an exact replication of Experiment 1 with the exception that participants did not receive any information or task before Round 2.

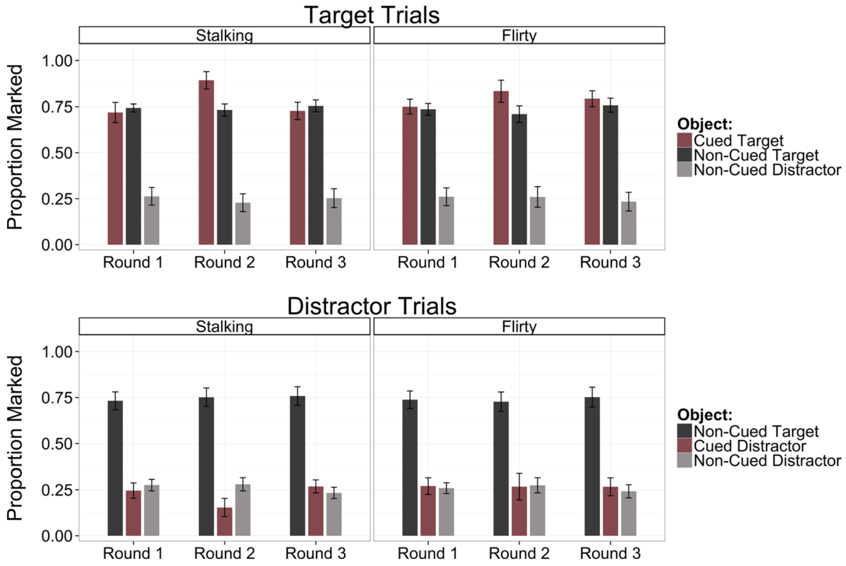

Fig. 4 Proportion marked of the different object types displayed in all three rounds for *cued-target* and *cued-distractor* trials depending on *cueing behavior*. Error bars indicate 95 % within-subjects confidence intervals.

Results

With an average of 26.05 correctly identified gaze expressions in the Reading- the-Mind-in-the-Eyes test, subjects in the present study scored minimally below the range of what is suggested as a normal control-group average by Baron-Cohen et al. (2001; 26.2 - 30.9 of 36 gaze expressions in total). No participant was excluded from the analysis.

Overall, participants tracked 2.87 of 4 objects correctly. In contrast to the previous experiments, there was a significant effect for *round*. Participants tracked more targets correctly in the second round, $t(19) = 3.11$, $p = .006.$, $d = 0.70$. The included contrasts were based on a linear mixed-effects model, fit via maximum likelihood, with *participants* as the random effect and the variables *condition* (Round 1, Round 2), *object status* (cued, noncued) and *cueing behavior* (flirty, stalking) as fixed effects. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .0125 per test (.05/4). Refer to Fig. 5 for a graphical display of the analyzed target and distractor trial data. See Appendix D for full results.

All planned comparisons for target trials were non-significant (see Appendix D). Results indicate that the selection preference of the participants was not influenced by gaze cues, neither in the first, nor in the second round of tracking: they did not mark the cued targets more often than the non-cued target. This was also independent of the presented *cueing behavior* (Flirty, Stalking). Equal to the target trials, the gaze cue did not give rise to a selection preference for the cued distractor. Whether the cue was provided as *flirty* or as *stalking* did not make a difference without an ACS.

Discussion

The third experiment was conducted to measure possible effects due to learning, that is, a possible interaction of tracking "experience" (Round 1, Round 2) and cue usability. None of the planned comparisons yield significant results, indicating that the participants did not process the gaze-cue as an object feature during tracking. The data reported are in agreement with previous findings stating that feature processing is disregarded during MOT when spatiotemporal information is available constantly (Papenmeier et al., 2013) and that

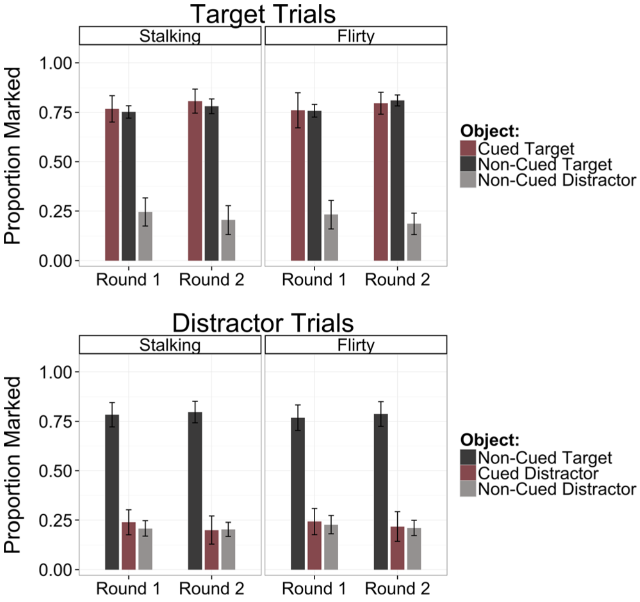

Fig. 5 Proportion marked of the different object types displayed in both rounds for *cued-target* and *cued-distractor* trials depending on *cueing behavior*. Error bars indicate 95 % within-subjects confidence intervals.

singletons (in this case, the cued object) are not processed differently when they are of no relevance to the given task (in this case, the tracking task) (e.g., Yantis & Egeth, 1994). We conclude that the use of features truly depends on our previously introduced manipulation (ACS). In the final experiment, we tested whether the ACS manipulation could be generalized to color cues by simply coloring neutral cartoon eyes for the length of a *Flirt* or the length of a trial (*Stalking*). The cued object remained white, while other objects were gray. We intended to replicate the results produced by the ACS effect.

Experiment 4

Method

Participants

Twenty new students of the University of Tübingen participated in return for monetary compensation (mean age M = 24.15 years, SD = 3.48). All subjects had reportedly normal or corrected-to-normal sight. Informed consent was obtained after the experimental procedures were explained to the subjects.

Stimuli and design

The design of the experiment was equal to the preceding Experiment 1. The participants tracked two rounds of 64 trials and received questions after each trial in the second round. They were also tested on their sensitivity to eyes (Read the Mind in the Eyes test) even though the stimuli used for the tracking task were changed to neutral eyes with motionless pupils (see Fig. 1). The participants were asked to track the *objects* that could, without further description, be perceived as resembling the figure 8 on its side (infinity sign) rather than cartoon eyes. Instead of taking a glance at the cued object, the noncued eyes were colored in a light gray tone while the cued object stayed white. In the *stalking* condition, noncued objects were colored during the whole trial (to be more specific, until 500 ms before the end of the motion), while in *flirty* trials the corresponding objects changed color twice, for the duration of 1.5 seconds (the exact duration of *flirt* moments in preceding experiments). The answer options were adapted (translated):

1. One of the **target**-objects was white; the other objects were gray.
2. One of the **distractor**-objects was white; the other objects were gray.
3. One of the **target**-objects changed its color to gray twice.
4. One of the **distractor**-objects changed its color to gray twice.

The participants received instructions with illustrations of the real tracking situations, similar to the instructions of Experiment 2 and 3.

Results

With an average of 26.75 correctly identified gaze expressions in the Reading the Mind in the Eyes test, subjects in the present study scored within the normal control-group range suggested by Baron-Cohen et al. (2001; 26.2–30.9 of 36 gaze expressions in total). The test was kept in the current experiment to ensure that differences do not arise due to a different course of the experiment.

Overall, participants tracked 2.85 of 4 objects correctly. We found a significant effect for *round*. Participants tracked more targets correctly in the second round, $t(19) = 2.18$, $p = .04$., $d = 0.49$.

Only correctly identified trials were included in the analysis, resulting in the inclusion of 70.9 % of the total trials of Round 2. The planned contrasts were based on a linear mixed-effects model, fit via maximum likelihood, with *participants* as the random effect and the variables *condition* (Round 1, Round 2), *object status* (cued, noncued) and *cueing behavior* (flirty, stalking). Refer to Fig. 6 for a graphical display of the analyzed target and distractor trial data. The according means and the results of the planned comparisons are presented in Appendix E. There were no significant differences between the selection of the cued-target objects and the noncued targets in Round 1, target trials *stalking,* $|z| = 0.03$, $p = .99$, $d = 0.01$, *flirty,* $|z| = 1.61$, $p = .43$, $d = 0.62$; distractor trials *stalking,* $|z| = 1.07$, $p = .77$, $d = 0.30$, *flirty,* $|z| = 1.04$, $p = .18$, $d = 0.58$. This reinsures that the color feature of the objects alone was not weighted as a source relevant enough to benefit tracking (in line with Papenmeier et al, 2013). Since color is known to be a preference for selective attention, this was especially remarkable with regard to the constantly visible white singleton among gray objects in the *stalking* trials. In Round 2, the interaction of ACS and the processing of color cues showed significant effects as seen before in Experiment 1, 2, and 3; target trials: *stalking,* $|z| = 5.39$, $p < .001$, $d = 1.11$, *flirty,* $|z| = 1.4.01$, $p < .001$, $d = 1.51$; distractor trials: *stalking,* $|z| = 4.59$, $p < .001$, $d = 1.47$, *flirty,* $|z| = 2.12$, $p = .15$, $d = 0.63$.

The results replicate what we have reported so far: The cued distractor was marked significantly less often. Whereas the cues on the target object in Round 2 worked independently of *cueing behavior*, the cued distractor was only marked less often than the rest of the distractors in the *stalking* condition compared to the *flirty* condition. Compared to the graphical inspection of the previous experiments, the current color experiment showed some slight indications for an automatic processing of abruptly occurring color cues in Round 1 (see tables in Appendix E and Fig. 6).

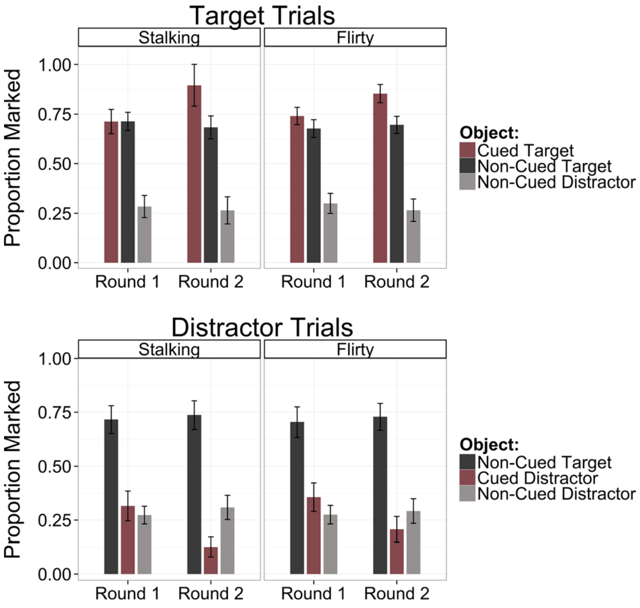

Fig. 6 Proportion marked of the different object types displayed in both rounds for *cued-target* and *cued-distractor* trials depending on *cueing behavior.* Error bars indicate 95 % within-subjects confidence intervals.

Discussion

The final experiment meant to generalize the effect found for gaze cues to color cues: Single cued target objects are marked more often than the other targets during tracking when the participant had an ACS (Round 2). Single cued distractor objects are marked less often, but this was only the case when the distractor was cued constantly compared to only twice. The results of Experiment 4 were remarkably similar to those of the previous experiments. They demonstrate that an ACS effect for single cued objects during a tracking task can be obtained with gaze cues as well as with color cues. The reasons why we shy away from claiming complete interchangeability between the two types of *cueing* (color and gaze) are the graphical results of Experiment 4. The results of the *flirty* trials in Round 1 hint to a different processing of color cues, which could have been due to its rather *abrupt* nature. As Yantis (1993) suggested, those kinds of visual onsets may capture attention independent of an attentional state of feature readiness.

A note on learning and task effects

We wondered about the effect of the identification task on general tracking performance. To that end we calculated the mean difference of proportion marked for the three noncued targets of Round 2 minus Round 1 (leaving out Round 3 of Experiment 2) independent of condition and behavior and applied two-sided *t* tests. The calculated differences were significantly different from zero in the first three experiments (see Fig. 7), Experiment 1: $t(19) = -2.38$, $p = 0.03$; Experiment 2: $t(31) = -2.58$, $p = 0.01$; Experiment 3: $t(19) = 2.58$, $p = 0.02$. The negative *t* values as well as the graphical presentation suggest that the identification task had a detrimental influence on the overall tracking performance in Experiment 1 and Experiment 2. In Experiment 3, when participants were not given the identification task in Round 2, they showed a (positive) learning effect. Interestingly though, Experiment 4 showed neither, $t(19) = 0.09$, $p = 0.93$. This could suggest that the task to identify a specific color "behavior" does not interfere with the tracking task at all. However, with regard to the learning effect in Experiment 3, it may well be assumed that the effect of learning is simply canceled out by the detrimental effect of the identification task.

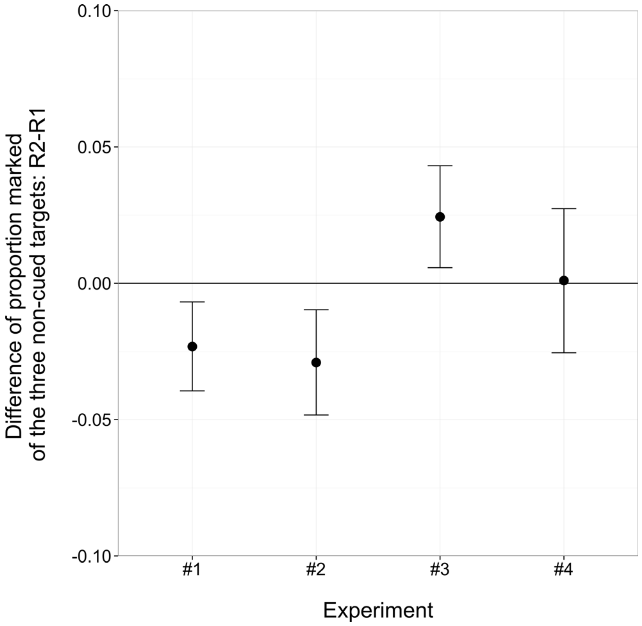

Fig. 7 Mean difference of proportion marked of the three noncued targets of (Round 2 minus Round 1) compared to zero (H_0: no difference between Round 1 and Round 2). Error bars indicate 95 % within-subjects confidence intervals.

General discussion

In the experiments presented, we explored how manipulating the relevance of object features modifies the allocation of attention in a MOT task. While it has been theorized that the attentional resolution is increased and decreased based on spatiotemporal occurrences like crowding or overlapping objects, we assume that an intentional, task-based component can influence attentional resolution to access feature objects as well. Our research design has a practical share with the concept of attentional capture and control settings (e.g., Egeth & Yantis, 1997; Folk et al., 1992) which lead us to hypothesize that changes in attentional resolution will happen in a strategic manner, inhibiting distractor objects and preferring target objects flexibly based on task-demands.

Theoretical accounts on tracking and features

The first interesting finding is reflected in the null-effects in the first round of each experiment: While it is well known that in the MOT paradigm it is easier to track the targets if they are easily distinguishable from the distractors, for example by being of a different color (Howe & Holcombe, 2012; Makovski & Jiang, 2009a; Makovski & Jiang, 2009b; similar: Oksama & Hyönä, 2008), our analysis of the selection preference for singletons and nonsingletons does not support this claim. However, Makovski and Jiang (2009a) also stated that features are not properly conjoined during attentive tracking because feature-location bindings are not necessary to track successfully. They concluded that the attentional tracking system has no access to bound target representations. Nonetheless, the model of multiple identity tracking (MOMIT; Oksama & Hyönä, 2008) predicts greater accuracy for distinctive targets because the observer can recover from tracking errors more easily. Based on the MOMIT, we would have expected that the target singleton is automatically (i.e. Round 1 of the experiments) marked more often, while the distractor singleton is confused less with a target than the other distractors—but that was not the case, not even in Experiment 4 with color singletons. This does not necessarily contradict previous findings, since for example Howe and Holcombe (2012) showed that the greatest advantage of distinct features is observed when distractor objects share absolutely none of the target features.

Our target and distractor objects shared the exact same features, that is, white spheres with pupils that moved along with the items and the singleton they cued (or as in Experiment 4: colored gray) with the exception that the singleton displayed "neutral" pupils (or stayed white). Our results suggest that focusing attention toward minimal featural differences during tracking is, first of all, possible, and more importantly, probably leads to a comparable advantage as if tracking targets and distractors that are different in *all* featural aspects.

If we disregard the difference between identity and feature and assume that the MOMIT (Oksama & Hyönä, 2004, 2008) applies to our stimuli, it would not explain our results entirely either. The MOMIT proposes that an *automatic*, parallel low-level system collaborates with a serial high-level attentional spotlight. When a target is at risk to fall below an activation threshold, the spotlight is alarmed by the low-level system, initiating an exhaustive search by the spotlight to reactivate the representation of the particular object in the low-level system. In order to explain our results, the model would need a modification: The spotlight is not only controlled by a low-level system but should also work in a top-down manner, even inhibiting stimuli with features that have no relevance.

The FLEX account by Alvarez and Franconeri (2007) describes a mechanism for the reallocation of resources in specific situations, for example, when a distractor comes close to a tracked target. They hypothesized that such interfering events would enhance the attentional resolution at the specific location. While we did not control for eye-movements in the current study, we found a way to approximate the proportion of attention each target and each distractor received, and observed attention shifts in favor of the identification task (see also Yarbus, 1967). That means, participants were able to find the singleton and/or follow the gaze cues (use the color cues) and used the information to avoid confusions with the cued distractor, or to have a "save" target (in the cued-target trials). However, whether attention was enhanced for targets and suppressed for distractors (as seen in Bettencourt & Somers, 2009, or Doran & Hoffman, 2010) cannot be determined with the current design. What we can propose is that, in principal, the shifted selection preferences may reflect attentional resolution. Observers can modulate the attentional resolution not only in accordance with spatial demands (as previously suggested by, e.g., Franconeri et al. 2008; Intriligator & Cavanagh, 2001; Pylyshyn, 2004) but also to access object feature information that based on task demands.

Underestimated objects: Distractors

As in Feria (2012), the current study not only emphasizes the role of distractors but also the role of distractor features during tracking. Especially interesting is the finding, that the singleton distractor is marked less often compared to other, noncued distractors in *stalking* conditions compared to the *flirty* conditions. In terms of existing tracking research, this observation does not support the majority of findings and could not reasonably be foreseen. In general, tracking capacity is greater when no other distracting objects are present (Horowitz & Cohen 2008; Horowitz et al. 2007) and decreases when the number of distractors increases (Bettencourt & Somers, 2009). Furthermore, by using a secondary probe detection task, researchers found that the detection of probes was less reliable on distractors compared to probes on targets and on an empty background (see Pylyshyn, 2006; Pylyshyn, Haladjian, King, & Reilly, 2008; Huff et al., 2012; but see Drew et al., 2009).

However, while these results point towards distractor inhibition during tracking, others reported that the role of distractors has been underestimated in its influence on tracking performance. Supporting findings of Alvarez and Oliva (2008), who showed that the locations of distractors are represented above chance level, Meyerhoff and team (2015) demonstrated that distractor displacements impair tracking performance. In a recent study, Meyerhoff, Papenmeier, Jahn, and Huff (in press) compared different speed profiles and reported an affected tracking performance even when only distractor objects varied in speed. They suggested that the spatial configuration of targets *and* distractors, in contrast to tracking for example only a virtual polygon of target objects (Yantis, 1992), is encoded in order to enhance the allocation of visual attention toward target objects. While they propose further that the distractor location is represented to an extent that allows for the detection of crowding events (see also Iordanescu et al., 2009), our data indicate that not only distractor *location* is represented but also a specified distractor *feature*.

In our data, the feature representation on the distractor object was at least pronounced enough to detect approaching events and to allocate attention accordingly in order to prevent confusions. Our findings nicely fit the results of Drew et al. (2009). In their tracking study with moving and stationary objects, they measured the electrophysiological responses (ERP) of participants to task-irrelevant probes that were located on targets, distractors, stationary objects, or in empty space. The authors report the response to distractors as located between the (greatest) response to targets and the (weakest) response to background and empty space probes. They concluded that distractors are not suppressed, at least not on an early level of perceptual processing. The idea of a hierarchy of attentional allocation (Drew et al., 2009)—with distractors being secondary to the dominant targets, but nonetheless processed—could help to explain why the singleton distractor is marked less often compared to other, noncued distractors in *stalking* conditions but not in *flirty* conditions. Given that the task triggered a feature level of attention, which changed the distribution of attention on the moving objects, distractors still received less attention than targets due to the tracking task demands. The intermittent *flirty* cue was simply too weak to be processed within the limited amount of attention allocated to a distractor, while the constantly visible *stalking* cue was stronger and thus needed less resources to have an attentional effect.

Intentional attentional control

However, it should be noted that we did not measure automatic effects that could be contributed to the featural singletons in Round 1. Features were only represented or at least measurably used during tracking in conditions in which we induced an ACS. Approaching the results in terms of attentional control and goal- related processing of singletons, it is easier to explain why targets were prioritized and distractor inhibited. Folk and team (1992) proposed that a task-driven selection mechanism guides attention and by that, the observer is able to ignore or prioritize distinct objects. Ignoring the distractor was harder to accomplish (i.e., demanded more cognitive resources) in the *flirty distractor* condition because participants had to keep track of four targets and figure out whether a cue that only appeared twice was beneficial or not. In the *stalking distractor* condition, the cue was constantly visible, and confusions with the cued distractor were less likely to occur. The same logic applies for tracking Round 1 (and Round 3 in Experiment 2). Without the task, no ACS was activated and thus cued single objects had no relevance. The spatiotemporal information was sufficient to track the objects successfully. This is what has been observed in previous studies and what has led to the conclusion that MOT is a feature-independent, preattentive and low-level task (e.g., Huff, Jahn, & Schwan, 2009; but see Papenmeier et al., 2013).

Costs of the additional task for the overall tracking performance

Concerning the overall tracking performance, we were certain that the additional identification task would draw upon cognitive resources, resulting in a lower performance in the dual-task conditions. However, based on our results of the control experiment compared to the experiments in which the manipulation was applied (Experiment 3 vs. 1, 2, and 4), we can only assume that the costs of the second task are basically rather small, and, in case of the Experiment 4 (color), congruent with small learning effects in tracking. Considering Cohen et al. (2011), who maintained that feature and location processing during tracking draw on the same, single cognitive resource—a claim further supported by neurophysiological and functional neuroimaging studies (e.g., Corbetta & Shulman, 2002; Sàenz, Buraĉas, & Boynton, 2003), that found brain regions for attention to feature and location to overlap—this is a surprising finding. In contrast to conclusions made by Cohen et al. (2011), the small costs of the additional identification task found in our data would rather support the notion that tracking is either handled by an entirely separate, encapsulated system from feature processing, or some sharing of resources is possible without much decrement. That is, when tracking a group of children of which one is your own offspring on the playground, you will be able to (attentionally) prioritize your own while at the same time your tracking ability for the others will not be interfered tremendously. Especially when the children's features, for example clothes, are colored differently.

Our findings do not necessarily contradict the notion that there is a trade-off between locations and identities completely. Our findings simply bring us a step further to the identification of the scope and limits of the involved resources, suggesting that the identity of only one object can be processed with negligible decrement to the overall location-based tracking performance of the other objects.

Closely related to Luck et al. (1996), who proposed that different processing levels control different types of attentional overload and interferences, Cohen et al. (2011) also presented evidence that mental resources can be voluntarily distributed across targets depending on task-demands (identity tracking or location tracking). The current study provides additional evidence that attention allocation during MOT has, or can be influenced by, a top-down component as well. This was reflected in the strategically suppression of distractor objects to avoid errors.

Flexible gaze-cue usability

A final word on our choice for the stimuli is needed. First and foremost, we used gazing eyes because we expected to observe reflexive attention shifts to the gazed-at object, even when the gaze cue was counterpredictive of the intended saccade direction (which would have been especially disturbing in cued-distractor trials; Kuhn & Kingstone, 2009). Gaze following is supposed to be reflexive and independent of cognitive load (e.g., Driver et al., 1999; Friesen & Kingstone, 1998). This equally applies for objects within dynamic displays, indicating that attention induced by such cues can be attached to moving objects and not only cue a spatial area (Marotta, Casagrande, & Lupiáñez, 2013).

Yet, while there is compelling evidence for a highly automatic behavior as a response to gaze cues, recent studies, including the presented one, raise some doubts (e.g., Koval, Thomas, & Everling, 2005). Numerous studies suggest that gaze cues can be used with a degree of flexibility (e.g., depending on the observers goal: Bayliss, Frischen, Fenske, & Tipper, 2007; Brooks & Meltzoff, 2005; Johnson, Slaughter, & Carey, 1998; Macdonald & Tatler, 2013; Ricciardelli, Carcagno, Vallar, & Bricolo, 2013). Particularly Böckler, Knoblich, and Sebanz (2011) use the gaze-cue paradigm to show that attention sharing (operationalized as mutual gaze (i.e., at least two pairs of eyes/faces shift their gazes simultaneously) can modulate joint attention. They not only propose that observing others sharing attention increases the significance of an ensuing gaze, but with a crucial condition in which they tested whether gaze following was modulated by the relevance of the looked-at-target to the observer's current goal/task, they concluded that certain contextual conditions and top-down mechanisms affect gaze-following behavior.

Our study provides further support for the possibility that gaze cues are not necessarily followed reflexively independent of cognitive load. Based on the non significant Round 1 results, we conclude that participants may have used the gaze-cue flexibly and strategically during tracking in our study. Still, it is difficult to determine whether they actually followed the gaze cue or concentrated on featural differences only. One finding that would speak in favor of gaze-cue processing is that the *stalking* and the *flirty* target trials showed the same effects. However, while this argues in favor of parallel processes one could also defend a serial account in combination with the hierarchy of attention as proposed by Drew et al. (2009) and discussed before. The participants would have scanned the targets first before scanning the distractors. The short duration of feature visibility in flirty trials (3 seconds in total) may have been not enough to process features and identify the behavior in distractor trials, since by the time the participant reached the first distractor object, the featural cue would have already disappeared.

On the other hand, the Experiment 4 (color cues) showed that the same attentional control applies to color stimuli—but still does little to clarify the specific issue of cue—versus feature use. While replicating the results of Experiments 1–3, in Experiment 4 with color cues we found some indication, even though not reaching statistical significance, of reactions to an abrupt onset of the color cue in *flirty* trials in Round 1. As Yantis (1993) suggested, those kinds of visual onsets may capture attention independent of an attentional state of feature readiness. Therefore, it is difficult to arrive at a definite conclusion as to whether gaze cues were used or object features were compared during tracking. We may have simply observed here that gaze cues were used intentionally but did not work reflexively, while color cues elicited bottom-up reactions that were actively suppressed and channeled in the ACS condition. Possible future research could be concerned with abrupt occurrences of salient features during tracking. By further connecting MOT to other fields of research (e.g., attention capture), we may be able to solve some of the riddles and misunderstandings that tracking studies could have not disentangle up to now.

Benefits, drawbacks, and further applications

The presented novel variation of the standard MOT task (Pylyshyn & Storm, 1988) kept the general structure of the task but presented a feature singleton among identical objects, that was either among the targets or among the distractors—a design that has not been applied before. This allowed us to explore various effects at the same time, which may simultaneously represent a benefit as well as a drawback. Here, we focused on activating an ACS and observed the prioritization or inhibition of the feature singleton. Nonetheless, additional beneficial information could have been gained from including a focus on the social aspect and/or including a clinical sample. We believe that our modification of the paradigm can be applied with clinical populations—for example, through testing the ability to switch between parallel and serial processing, or testing patients with autism or Asperger's on their processing skills of dynamic gazes. With this being said, an eye-tracking study will be of tremendous use to further understand the modified paradigm. One could determine whether color and gaze cues produce the same results but are processed differently. In other words, gaze cues may produce involuntary saccades to the cued object but could be actively suppressed by the observer in order to successfully track the objects. Furthermore, we may find parallel processing (focus on the centroid of the targets) in Round 1, and serial processing (target jumping) in Round 2. Another option could be the use of single-pulse transcranial magnetic stimulation (TMS) on the superior lateral temporal cortex that is known to interfere with gaze direction tasks (Pourtois, Grandjean, Sander, & Vuilleumier, 2004). This interference was also found to be task-specific.

Although we did not find direct evidence for an automatic distinctiveness effect for the feature singletons, it is possible that we failed to measure the effect due to our experimental design. Future experiments concerned with automatic attentional shifts in MOT environments should consider contrasting trials with and without singletons, that is, including trials without distinct objects. Furthermore, the slightly different results found in Experiment 4 (color), however not statistically significant, could be an indicator of reflexive attention shifts in case of abrupt appearing cues. Future studies could be concerned with bottom-up effects of abrupt and gradual appearing features in MOT studies, and its dependency on an ACS.

Finally, we would like to highlight our decision for adapting the standard analysis of tracking capacity to our hypotheses. In contrast to the majority, if not all, MOT studies, we measured how often each of the eight objects was marked and compared this selection mean of cued and noncued objects by condition. This gave us a far more vivid picture of the attention distribution than a simple mean value of correctly tracked items. In fact, proportion correct showed no effects in our data (i.e., was not affected by the *cueing behavior* displayed). We believe this approach to be promising for future research (e.g., for studies concerned with the different role of targets and distractors). It is our hope that our newly modified analysis will be in use in future studies in order to gain more refined insight. However, it is important to mention that our decision to include only trials in which the participants correctly identified the eyes behavior and the type of object that was cued, was a theoretical advantage (excluding trials in which we cannot be sure that the participants was actively involved in the task), but a slight disadvantage for interpretative purposes. Because the proportion of correctly identified trials differed for the two gazing behaviors, we only analyzed data within each behavior and each round, and not across behaviors and rounds. Any statements concerned with the strength of the effect found in *flirty* compared to *stalking* conditions are thus purely speculative. Future studies focusing on ACS and MOT should find a way to control the participant's attentional engagement with less consequences for data analysis. Regardless, the present results indicate a consistent pattern of attentional resolution in tracking tasks that changes due to task demands.

Conclusion

Our results reveal a striking cued-target selection preference and a cued- distractor inhibition when participants received an identification task that engaged them actively in the processing of object features during tracking. These effects were attributed to the activation of an *attentional control set,* a term that originally stems from research on attention capture and visual search. We propose here that the allocation of attention and a flexible attentional resolution is not only an automatic reaction of the visual system to prevent confusions when interobject spacing decreases but also managed by a goal-related, top-down component. The introduced modification of the MOT paradigm, as well as the unusual type of analysis, offer various new options for future research in different areas.

Author Note Alisa Brockhoff & Markus Huff, Department of Psychology, University of Tübingen.

We thank Dominik Frommeyer and Lia Heubner for conducting the experiments. We are also grateful to Julian Hofmeister and Steffen Hildebrandt from the computer science department. We thank Linda Kersten for proofreading and Annika Maurer for general support. We sincerely thank Trafton Drew and two anonymous reviewers for constructive criticisms and valuable comments that have helped tremendously in the improvement of the overall presentation and clarity of this work.

This research was supported by grants from the German Research Foundation (DFG 1510/4-3). It has been partially presented (poster) at the annual meeting of the Vision Science Society in Florida, 2015, and at the Psychonomic Society's 56th Annual Meeting held in Chicago, 2015.

Appendix A

Table 2 Chi-square tests for differences in flirty and stalking identifications with Yate's continuity correction for small samples:

Flirty vs. stalking	X^2	p	CI (%)	Difference (%)
Experiment 1	31.15	>.001	[.07; .14]	9.56
Experiment 2	6.43	.001	[.01; .10]	5.31
Experiment 4	9.46	.002	[.04; .16]	7.94

Appendix B

Table 3 Experiment 1 Proportion marked of the different object types (cued, noncued) in the two conditions (R1, R2) and the two cueing behaviors (F, S) and results of the planned simultaneous comparisons for cued target trials.

M (*SD*)			$\lvert\text{Estimate}\rvert$(*SD*)	$\lvert z\rvert$	*p*
R1, S, CT	.74 (.15)	H_1]	.02 (.03)	.697	.928
R1, S, NCT	.73 (.10)				
R1, F, CT	.72 (.14)	H_2]	.00 (.03)	.223	.999
R1, F, NCT	.72 (.10)				
R2, S, CT	.87 (.19)	H_3]	.15 (.03)	5.40	<.001
R2, S, NCT	.71 (.11)				
R2, F, CT	.85 (.12)	H_4]	.17 (.03)	6.14	<.001
R2, F, NCT	.68 (.13)				
R2, S, CT - NCT vs. R2, F, CT - NCT		H_5]	02 (.04)	.524	.973

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .01 per test (.05/5). R1 = Round 1, R2 = Round 2, F = flirty, S= stalking, CT= cued target, NCT = non-cued targets.

Table 4 Experiment 1 Proportion marked of the different object types in the two conditions (R1, R2) and the two cueing behaviors (F, S) and results of the planned simultaneous comparisons for cued distractor trials.

	M (*SD*)		$\lvert\text{Estimate}\rvert$ (*SD*)	$\lvert z\rvert$	*p*
R1, S, CD	.29 (.13)	H_1]	.03 (.03)	1.04	.774
R1, S, NCD	.26 (.10)				
R1, F, CD	.26 (.13)	H_2]	.01 (.03)	.351	.995
R1, F, NCD	.27 (.11)				
R2, S, CD	.23 (.18)	H_3]	.03 (.03)	2.26	.101
R2, S, NCD	.29 (.11)				
R2, F, CD	.28 (.13)	H_4]	.01 (.03)	.212	.999
R2, F, NCD	.29 (.11)				
R2, S, CT - NCT vs. R2, F, CT - NCT		H_5]	.06 (.04)	1.45	.491

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .01 per test (.05/5). R1 = Round 1, R2 = Round 2, F = flirty, S = stalking, CD = cued distractor, NCD = non-cued distractors.

Appendix C

Table 5 Experiment 2 Proportion marked of the different object types (cued, noncued) in the three conditions (R1, R2, R3) and cueing behavior (stalking, flirty) and results of the planned simultaneous comparisons for cued target trials.

	M (*SD*)		$\lvert\text{Estimate}\rvert$(*SD*)	$\lvert z\rvert$	*p*
R1, S, CT	.72 (.17)	H_1]	.02 (.03)	.767	.974
R1, S, NCT	.74 (.09)				
R1, F, CT	.75 (.13)	H_2]	.01 (.03)	.441	.999
R1, F, NCT	.74 (.11)				
R2, S, CT	.89 (.15)	H_3]	.16 (.03)	5.08	< .001
R2, S, NCT	.73 (.11)				
R2, F, CT	.83 (.17)	H_4]	.12 (.03)	3.873	<.001
R2, F, NCT	.71 (.14)				
R3, S, CT	.73 (.15)	H_5]	.03 (.03)	.847	.958
R3, S, NCT	.75 (.11)				
R3, F, CT	.79 (.14)	H_6]	.03 (.03)	1.10	.870
R3, F, NCT	.76 (.12)				
R2, S, CT - NCT vs. R2, F, CT - NCT		H_7]	.28 (.04)	6.32	<.001
R3, S, CT - NCT vs. R3, F, CT - NCT		H_8]	.01 (.04)	.178	.999

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .00625 per test (.05/8). R1 = Round 1, R2 = Round 2, R3 = Round 3, S = stalking, F = flirty, CT = cued target, NCT = non-cued targets.

Table 6 Experiment 2 Proportion marked of the different object types (cued, noncued) in the three conditions (R1, R2, R3) and cueing behavior (stalking, flirty) and results of the planned simultaneous comparisons for cued-distractor trials.

	M (*SD*)		|Estimate|(*SD*)	|*z*|	*p*
R1, S, CD	.24 (.14)	H_1]	.03(.03)	.926	.936
R1, S, NCD	.28 (.11)				
R1, F, CD	.27 (.15)	H_2]	.01 (.03)	.347	.999
R1, F, NCD	.26 (.10)				
R2, S, CD	.15 (.15)	H_3]	.13 (.03)	3.81	<.001
R2, S, NCD	.28 (.11)				
R2, F, CD	.27 (.23)	H_4]	.01 (.03)	.225	.999
R2, F, NCD	.27 (.12)				
R3, S, CD	.27 (.12)	H_5]	.04 (.03)	1.12	.858
R3, S, NCD	.23 (.12)				
R3, F, CD	.26 (.16)	H_6]	.02 (.03)	.775	.973
R3, F, NCD	.23 (.12)				
R2, S, CD - NCD vs. R2, F, CD - NCD		H_7]	.13 (.05)	2.90	<.003
R3, S, CD - NCD vs. R3, F, CD - NCD		H_8]	.06 (.04)	1.34	.725

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .00625 per test (.05/8). R1 = Round 1, R2 = Round 2, R3 = Round 3, S = stalking, F = flirty, CT = cued distractor, NCT = non-cued distractors.

Appendix D

Table 7 Experiment 3 Proportion marked of the different object types in the two conditions (R1, R2) and the two cueing behaviors (F, S) and results of the planned simultaneous comparisons for cued target trials.

	M (*SD*)		|Estimate| (*SD*)	|*z*|	*p*
R1, S, CT	.76 (.16)	H_1]	.02 (0.04)	.389	.992
R1, S, NCT	.75 (.08)				
R1, F, CT	.76 (.22)	H_2]	.00 (0.04)	.060	1.00
R1, F, NCT	.76 (.09)				
R2, S, CT	.81 (.16)	H_3]	.03 (0.04)	.658	.943
R2, S, NCT	.78 (.11)				
R2, F, CT	.80 (.15)	H_4]	.01 (0.04)	.359	.994
R2, F, NCT	.81 (.07)				

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .0125 per test (.05/4). R1 = Round 1, R2 = Round 2, F = flirty, S = stalking, CT = cued target, NCT = non-cued targets.

Table 8 Experiment 3 Proportion marked of the different object types in the two conditions (R1, R2) and the two cueing behaviors (F, S) and results of the planned simultaneous comparisons for cued distractor trials.

	M (*SD*)		|Estimate| (*SD*)	|*z*|	*p*
R1, S, CD	.24 (.15)	H_1]	.03 (.04)	.798	.891
R1, S, NCD	.21 (.10)				
R1, F, CD	.24 (.14)	H_2]	.02 (.04)	.399	.991
R1, F, NCD	.23 (.11)				
R2, S, CD	.20 (.16)	H_3]	.00 (.04)	.092	1.00
R2, S, NCD	.20 (.09)				
R2, F, CD	.22 (.18)	H_4]	.01 (.04)	.184	1.00
R2, F, NCD	.21 (.10)				

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .0125 per test (.05/4). R1 = Round 1, R2 = Round 2, F = flirty, S = stalking, CD = cued distractor, NCD = non-cued distractors.

Appendix E

Table 9 Experiment 4 Proportion marked of the different object types (cued, non-cued) in the two conditions (R1, R2) and the two object shapes (E, A) and results of the planned simultaneous comparisons for cued target trials.

	M (*SD*)		|Estimate|(*SD*)	|*z*|	*p*
R1, S, CT	.71 (.14)	H_1]	.00 (.04)	.026	1.00
R1, S, NCT	.71 (.10)				
R1, F, CT	.74 (.11)	H_2]	.06 (.04)	1.61	.43
R1, F, NCT	.67 (.10)				
R2, S, CT	.90 (.23)	H_3]	.21 (.04)	5.39	<.001
R2, S, NCT	.68 (.13)				
R2, F, CT	.86 (.18)	H_4]	.16 (.04)	4.01	<.001
R2, F, CT	.85 (.11)				
R2, E, CT - NCT vs. R2, A, CT - NCT		H_5]	.03 (.06)	.519	.990

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .01 per test (.05/5). R1 = Round 1, R2 = Round 2, S = stalking, F = flirty, CT = cued target, NCT = non-cued targets.

Table 10 Experiment 4 Proportion marked of the different object types in the two conditions (R1, R2) and the two cueing behaviors (F, S) and results of the planned simultaneous comparisons for cued distractor trials.

	M (*SD*)		|Estimate| (*SD*)	|*z*|	*p*
R1, S, CD	.32 (.17)	H_1]	.04 (.04)	1.07	.772
R1, S, NCD	.27 (.10)				
R1, F, CD	.35 (.16)	H_2]	.08 (.04)	1.036	.177
R1, F, NCD	.28 (.11)				
R2, S, CD	.13 (.12)	H_3]	.18 (.04)	4.59	<.001
R2, S, NCD	.30 (.13)				
R2, F, CD	.21 (.15)	H_4]	.08 (.04)	2.12	.146
R2, F, NCD	.29 (.13)				
R2, E, CT - NCT vs. R2, A, CT - NCT		H_5]	.25 (.06)	4.45	<.001

Note. A priori hypotheses were conducted using Bonferroni adjusted alpha levels of .01 per test (.05/5). R1 = Round 1, R2 = Round 2, E = eyes, A = arrows, CD = cued distractor, NCD = non-cued distractors.

References

Adams, R. B., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14*(6), 644–647.

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

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392–398.

Bae, G. Y., & Flombaum, J. I. (2012). Close encounters of the distracting kind: Identifying the cause of visual tracking errors. *Attention, Perception, & Psychophysics, 74*(4), 703–715.

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological bulletin, 112*(3), 461.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*, 395–418.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes?. *British Journal of Developmental Psychology, 13*(4), 379–398.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241–251.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition, 104*(3), 644–653.

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.

Böckler, A., Knoblich, G., & Sebanz, N. (2011). Observing shared attention modulates gaze following. *Cognition, 120*(2), 292–298.

Bölte, S. (2005). *Reading Mind in the Eyes Test für Erwachsene (dt. Fassung) von S. Baron-Cohen [Reading the Mind in the Eyes test for adults (German version) by S. Baron-Cohen], JW Goethe Universität Frankfurt am Main: Eigendruck.*

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535–543.

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

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review, 6*(2), 204–223.

Cohen, M. A., Pinto, Y., Howe, P. D., & Horowitz, T. S. (2011). The what-where trade-off in multiple-identity tracking. *Attention, Perception, & Psychophysics, 73*(5), 1422–1434.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Cutzu, F., & Tsotsos, J. K. (2003). The selective tuning model of attention: Psychophysical evidence for a suppressive annulus around an attended item. *Vision Research, 43*(2), 205–219.

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.

Drew, T., McCollough, A. W., Horowitz, T. S., & Vogel, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review, 16*(2), 411–417.

Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition, 126*(2), 213–223.

Driver, J., IV, Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*(5), 509–540.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*(1), 269–297.

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*(4), 225–240.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5*(1), 39–60.

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.

Folk, C. L., & Annett, S. (1994). Do locally defined feature discontinuities capture attention? *Perception & Psychophysics, 56*(3), 277–287.

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 and Performance, 18*(4), 1030.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 317.

Franconeri, S. L., Lin, J. Y., Enns, J. T., Pylyshyn, Z. W., & Fisher, B. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic bulletin & review, 15*(4), 802–808.

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*(7), 920–925.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*(3), 490–495.

Hillstrom, A. P., & Yantis, S. (1994). Visual motion and attentional capture. *Perception & Psychophysics, 55*(4), 399–411.

Horowitz, T. S., & Cohen, M. A. (2008). How do distractors distract in multiple object tracking? *Perception, 37,* 98–99.

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.

Howe, P.D., & Holcombe, A.O. (2012). The effect of visual distinctiveness on multiple object tracking performance. *Frontiers in Psychology, 3*(307), 1–7. doi:10.3389/fpsyg.2012.00307

Huff, M., Jahn, G., & Schwan, S. (2009). Tracking multiple objects across abrupt viewpoint changes. *Visual Cognition, 17*(3), 297–306.

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

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*(3), 171–216.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of velocities during multiple-object tracking. *Journal of Vision, 9*(4), 1.

James, W. (1890). A saliency-based search mechanism for overt and covert shifts of visual attention. *Principles of Psychology, 2, Harvard University Press, Cambridge, MA (1890/1980).*

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1*(2), 233–238.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43*(4), 346–354.

Joseph, J. S., & Optican, L. M. (1996). Involuntary attentional shifts due to orientation differences. *Perception & Psychophysics, 58*(5), 651–665.

Koval, M. J., Thomas, B. S., & Everling, S. (2005). Task-dependent effects of social attention on saccadic reaction times. *Experimental Brain Research, 167*(3), 475–480.

Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception, & Psychophysics, 71*(2), 314–327.

Laberge, D., & Brown, V. (1986). Variations in size of the visual field in which targets are presented: An attentional range effect. *Perception & Psychophysics, 40*(3), 188–200.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*(9), 1183–1198.

Luck, S. J., Hillyard, S. A., Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception and Performance, 22*(3), 725.

Lupiáñez, J., Milliken, B., Solano, C., Weaver, B., & Tipper, S. P. (2001). On the strategic modulation of the time course of facilitation and inhibition of return. *The Quarterly Journal of Experimental Psychology: Section A, 54*(3), 753–773.

Macdonald, R. G., & Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of Vision, 13*(4), 6.

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

Makovski, T., & Jiang, Y. V. (2009b). The role of visual working memory in attentive tracking of unique objects. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1687.

Marotta, A., Casagrande, M., & Lupiáñez, J. (2013). Object-based attentional effects in response to eye-gaze and arrow cues. *Acta Psychologica, 143*(3), 317–321.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (Vol. 1). New York, NY: Psychology Press.

Meyerhoff, H.S., Papenmeier, F., Jahn, G., & Huff, M. (2015). Distractor locations influence multiple object tracking beyond interobject spacing. *Experimental Psychology, 62*, 170–180. doi:10.1027/1618-3169/a000283

Meyerhoff, H. S., Papenmeier, F., Jahn, G., & Huff, M. (in press). Not FLEXible enough: Exploring the temporal dynamics of attentional reallocations with the multiple object tracking paradigm. *Journal of Experimental Psychology: Human Perception and Performance.* doi:10.1037/xhp0000187

Morgan, M. J., Ward, R. M., & Castet, E. (1998). Visual search for a tilted target: Tests of spatial uncertainty models. *The Quarterly Journal of Experimental Psychology: Section A, 51*(2), 347–370.

Morgan, M. J., & Watt, R. J. (1997). The combination of filters in early spatial vision: A retrospective analysis of the MIRAGE model. *Perception-London, 26,* 1073–1088.

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology, 20*(01), 79–97.

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition, 11*(5), 631–671.

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.

Papenmeier, F., Meyerhoff, H.S., Jahn, G., & Huff, M. (2013). Tracking by location and features: Object correspondence across spatiotemporal discontinuities during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 159.

Peirce, J.W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*(10). doi:10.3389/neuro.11.010.2008

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes, 32*, 531–556.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. *Modes of Perceiving and Processing Information, 137,* 158.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*(6), 619–633.

Pratt, J., Radulescu, P. V., Guo, R. M., & Abrams, R. A. (2010). It's alive!: Animate motion captures visual attention. *Psychological Science, 21*(11), 1724–1730.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*(1), 65–97.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(03), 341–365.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition, 11*(7), 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. W. (2007). *Things and places: How the mind connects with the world*. Cambridge, MA: MIT Press.

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

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

Ricciardelli, P., Carcagno, S., Vallar, G., & Bricolo, E. (2013). Is gaze following purely reflexive or goal-directed instead? Revisiting the automaticity of orienting attention by gaze cues. *Experimental Brain Research, 224*(1), 93–106.

Ruxton, G. D., & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral Ecology, 19*(3), 690–693.

Sàenz, M., Buracâs, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43*(6), 629–637.

Scaife, M., & Bruner, J.S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266.

Theeuwes, J., Kramer, A. F., & Atchley, P. (1999). Attentional effects on preattentive vision: spatial precues affect the detection of simple features. *Journal of Experimental Psychology: Human Perception and Performance, 25*(2), 341.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological review, 95*(1), 15.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 459.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2005). Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience, 17*(12), 1907–1922.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*(3), 295–340.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 676–681.

Yantis, S., & Egeth, H. E. (1994). Visual salience and stimulus-driven attentional capture. *Investigative Ophthalmology & Visual Science, 35*(4), 1619–1619.

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York, NY: Plenum Press.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*(6706), 72–75.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research, 39*(2), 293–306.