

# Spatial updating of dynamic scenes: Tracking multiple invisible objects across viewpoint changes

MARKUS HUFF, HAUKE S. MEYERHOFF, AND FRANK PAPANMEIER  
Knowledge Media Research Center, Tübingen, Germany

AND

GEORG JAHN  
University of Greifswald, Greifswald, Germany

Research on dynamic attention has shown that visual tracking is possible even if the observer's viewpoint on the scene holding the moving objects changes. In contrast to smooth viewpoint changes, abrupt changes typically impair tracking performance. The lack of continuous information about scene motion, resulting from abrupt changes, seems to be the critical variable. However, hard onsets of objects after abrupt scene motion could explain the impairment as well. We report three experiments employing object invisibility during smooth and abrupt viewpoint changes to examine the influence of scene information on visual tracking, while equalizing hard onsets of moving objects after the viewpoint change. Smooth viewpoint changes provided continuous information about scene motion, which supported the tracking of temporarily invisible objects. However, abrupt and, therefore, discontinuous viewpoint changes strongly impaired tracking performance. Object locations retained with respect to a reference frame can account for the attentional tracking that follows invisible objects through continuous scene motion.

Interacting with a dynamic environment sometimes requires keeping track of several moving objects in parallel. In road traffic, for example, it is necessary to pay attention to cars, bikers, and pedestrians. However, not every object that is attended to in traffic is visible all the time; some are occluded for a short time (e.g., by a large truck driving by). Usually, individuals can handle such situations. They succeed in attentional tracking even if relevant participants in traffic are temporarily invisible. Moreover, observers in traffic situations often are themselves moving. Observing a dynamic scene while one is moving adds to the complexity of scene motion and, thus, the need to track objects in a scene that undergoes continuous viewpoint changes. How do individuals manage this and similar tasks?

Studies of multiple-object tracking (MOT) have shown that humans are able to track several independently moving targets among identical-looking distractors (Pylyshyn & Storm, 1988). In MOT studies, the objects to be tracked usually move in a static frame. Hence, there is object motion but no scene motion. Recently, however, MOT was examined with scenes in motion. Liu et al. (2005) showed that smooth scene motion—such as rotation, translation, and zoom—did not impair tracking performance. *Continuous* visual information about scene motion was sufficient to establish correspondence between targets from

one moment to the next. *Discontinuous* scene motion disrupted this moment-to-moment tracking. For example, abrupt rotational viewpoint changes strongly impaired tracking performance (Huff, Jahn, & Schwan, 2009; see also Seiffert, 2005). Thus, multiple objects can be tracked even if they move within a scene that is itself in motion, provided that there is continuous visual information about scene motion.

Abrupt, but not smooth, scene motion creates hard onsets of targets and distractors at their new locations. With smooth scene motion, there are no hard onsets at new locations. Such hard onsets could impair tracking performance, due to attentional capture (Pylyshyn, 2007). According to the FINST-theory fingers of instantiation (FINST) theory (Pylyshyn, 2001), preattentive indices grab suddenly appearing objects. Hard onsets could remove indices from targets and assign them to a novel set of objects. Thus, lower tracking performance with abrupt scene motion could be attributed to hard onsets, instead of to a lack of continuous scene information. To solve this confound, we can introduce short intervals of object invisibility during abrupt and smooth viewpoint changes. Then objects reappearing at their new locations create hard onsets with both abrupt and smooth viewpoint changes.

---

M. Huff, m.huff@iwm-kmrc.de

---

Without scene motion, MOT is robust against brief intervals of target invisibility. In several studies, tracking performance was hardly impaired if all objects simultaneously disappeared for several hundred milliseconds (Alvarez, Horowitz, Arsenio, Dimase, & Wolfe, 2005; Horowitz, Birnkrant, Fencsik, Tran, & Wolfe, 2006; Keane & Pylyshyn, 2006). Tracking performance depended on the distance between the locations at which objects disappeared and reappeared: the larger this distance, the lower the observed tracking performance (Keane & Pylyshyn, 2006). This finding suggests that target locations were briefly retained after targets had turned invisible. Keane and Pylyshyn postulated preattentive indices (FINSTs) that stick with the targets and persist for several hundred milliseconds after target disappearance. When the targets reappear, these FINSTs (Pylyshyn, 2001, 2007) grab an object within a specified region. That preattentive indices are the mechanism behind MOT is controversial. Instead, attentional foci that are kept on targets may underlie successful tracking (Cavanagh & Alvarez, 2005). Such foci may persist and lead to recollection of temporarily invisible targets as well. As a third alternative, it has been suggested that spatial memory for target locations explains MOT of temporarily invisible objects. In one study, tracking performance was better if all objects disappeared simultaneously, as compared with asynchronous vanishing (Horowitz et al., 2006). The authors suggested an *offline storage* that is triggered by cues from objects that vanish simultaneously and contains information about target positions at the moment right before disappearance. The offline stored target locations are assumed to be used to recollect reappearing targets.

Recent studies have provided evidence for object-based attentional phenomena that persist while objects are temporarily invisible. Flombaum, Scholl, and Pylyshyn (2008) showed that objects that moved behind occluding strips in MOT receive object-based attention. In their experiments, participants had to detect small probes on targets, distractors, occluders, or the empty space between them. Detection performance for probes on occluded objects (targets and distractors) increased, suggesting sustained object-based attention during intervals of invisibility (see also Scholl & Pylyshyn, 1999).

Attentional tags marking recently inspected objects can persist as well. The *inhibition-of-return* phenomenon is shown by slower probe detection at recently attended locations (Posner & Cohen, 1984; see also Klein, 2000). Object-based inhibition of return holds for moving objects, too (Tipper, Driver, & Weaver, 1991). Furthermore, invisibility of objects at any stage of a trial—during cuing, probing, or both—did not eliminate object-based inhibition of return (Yi, Kim, & Chun, 2003; see also Jefferies, Wright, & Di Lollo, 2005).

Without scene motion, the mechanisms retaining the locations of temporarily invisible objects seem to be sufficient; preattentive indices (Pylyshyn, 2001, 2007), attentional foci (Cavanagh & Alvarez, 2005), or target locations stored offline (Horowitz et al., 2006) could explain why tracking performance decreases with increasing distance between the locations of disappearance and reappearance.

However, if the scene holding the moving objects undergoes motion, statically retained target locations can easily lead to false identifications and, consequently, to impaired MOT performance. In order to keep track of objects that are invisible during scene motion, an updating process is essential. The invisible objects' retained locations have to be adapted in line with the objects' displacements resulting from scene motion. The updating process that we suggest is simple: If the retained locations were linked to visible scene elements that acted as a reference frame, they would follow scene motion. For this to work, locations of temporarily invisible objects have to be retained *online*.

Given that target locations are retained for temporarily invisible targets and that they may be linked to a visible reference frame, for the locations to be updated with scene motion, presumably, the information about scene motion has to be continuous. Discontinuous scene motion disrupts MOT for visible objects; hence, it should also disrupt MOT for objects invisible during scene motion. We were less sure about what we would observe with continuous scene motion. Therefore, the main question that we addressed in the experiments reported here is whether MOT is possible if objects are invisible during continuous scene motion.

### Experimental Overview

In all three experiments, observers tracked multiple targets in 3-D scenes that underwent either an abrupt or a smooth viewpoint change of 30°. At the time of the viewpoint change, the objects were either visible or invisible. There were also control conditions without viewpoint changes, in which objects were always visible or temporarily invisible.

In conditions with objects visible throughout the entire trial, we expected lower visual tracking performance with abrupt viewpoint changes, as compared with the conditions with smooth and without viewpoint changes (Huff et al., 2009; Liu et al., 2005). In conditions without viewpoint changes, we expected slightly impaired MOT performance if the objects were temporarily invisible (Keane & Pylyshyn, 2006). Because object translations during object invisibility are larger in conditions with viewpoint changes than in those without viewpoint changes, we expected lower visual-tracking performance in conditions with viewpoint changes. Finally, if invisible objects' locations are retained and linked to a reference frame, so that they are updated following continuous scene motion, performance with invisible objects should be higher with smooth than with abrupt viewpoint changes. If there is no updating process, tracking performance in these two conditions should be equal. The three experiments differed with regard to the amount of visual detail in the scenes and the number of the to-be-tracked target objects. In Experiment 1, targets and distractors moved on a floor plane with a checkerboard pattern; in Experiment 2, the floor plane was monochromatic. In these experiments, participants tracked two, three, or four targets. Finally, in Experiment 3, we directly compared the two floor planes in Experiments 1 and 2 at a constant tracking load of three target objects.

## EXPERIMENT 1

### Method

**Participants.** Twenty students who were recruited from the psychology department of the University of Tübingen (19 of them female) participated in the experiment. Their mean age was 21.8 years (19–25 years). All the participants reported normal or corrected-to-normal vision. They received compensation for their participation.

**Apparatus, Stimuli, and Procedure.** The stimuli were presented on an 18-in. LCD monitor ( $1,024 \times 768$  pixels) rendered online by algorithms written in Python, using the 3-D graphics software Blender. The participants were seated at a viewing distance of 55 cm, resulting in a viewable area of  $36.1^\circ$  width and  $29.3^\circ$  height.

The stimuli were 12 white spheres, which moved on a rectangular floor plane with a checkerboard pattern consisting of light gray and dark gray squares (see Figure 1). The floor plane appeared on a dark blue background. All the visible objects were drawn as if seen from a camera angle of  $20^\circ$  to the  $x$ - $y$  plane. The spheres subtended  $0.7^\circ$ – $1.3^\circ$  of visual angle, with an average of  $0.9^\circ$ , whereas the floor rectangle occupied  $15.2^\circ$ – $26.5^\circ$  horizontally and  $7.2^\circ$  vertically. At the beginning of each trial, the empty floor plane appeared. After 2 sec, the 12 spheres were positioned randomly on the floor plane, and 2, 3, or 4 of them were designated as targets by their flashing on and off in red four times in 200-msec intervals and remaining red for 2 sec before they turned back to white.

Then all the objects started to move in randomly selected directions with a constant speed of  $2^\circ/\text{sec}$ , bouncing off the edges of the floor rectangle in the manner of billiard balls. The participants were instructed to keep track of the previously designated targets throughout the trial. The moving interval was 11 sec, after which all the objects stopped and the participants responded by marking the spheres with mouse clicks. The marked spheres turned red and could be unmarked again. After marking as many spheres as there were targets on the current trial, the participants received feedback about the number of correctly identified targets.

We presented equal numbers of trials without, with smooth, and with abrupt viewpoint changes. Viewpoint changes had an extent of  $30^\circ$  clockwise or counterclockwise and occurred either abruptly at 8 sec into the movement interval or in a smooth way, which resembled a camera movement lasting 250 msec and starting at 7,875 msec. On half of the trials, all the targets and distractors disappeared instantaneously for 500 msec within 7,750 and 8,250 msec after movement onset. Examples of all the conditions are available at [www.iwm-kmrc.de/cybermedia/invisible-objects](http://www.iwm-kmrc.de/cybermedia/invisible-objects).

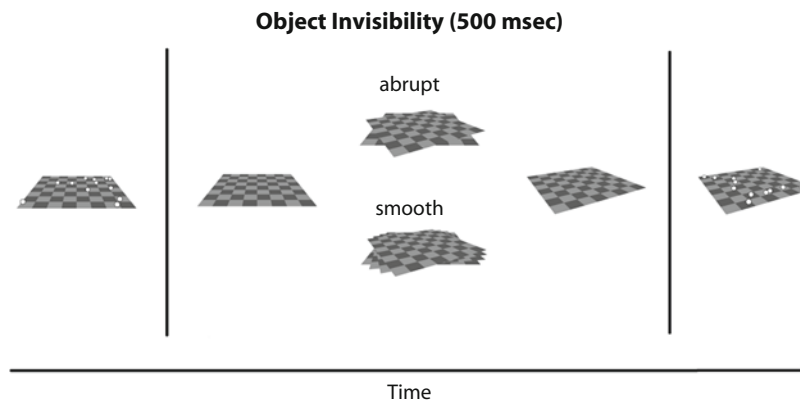
The variables *viewpoint change* (none, smooth, or abrupt), *visibility* (always visible or invisible for 500 msec), and *number of targets* (two, three, or four) were all varied within subjects. All the participants started with a training session of 18 practice trials. There were a total of 108 experimental trials, which were presented in random order.

### Results and Discussion

Proportion correct scores were calculated as the ratio of the number of correctly identified target objects on a trial to the total number of targets on that trial. Chance performance was characterized as the ability to track a single target object (Hulleman, 2005), which, accordingly, was .55, .45, and .45, for two, three, and four target trials, respectively. Mean proportions of correctly identified targets are plotted in Figure 2 separately for conditions with two, three, and four targets. All the plots show that tracking performance was better with smooth than with abrupt viewpoint changes, even if the objects were invisible during the viewpoint change. Proportion correct scores for all the participants were submitted to a  $3$  (viewpoint change)  $\times$   $2$  (visibility)  $\times$   $3$  (number of targets) factorial repeated measures ANOVA. We report partial eta-squared ( $\eta_p^2$ ) as effect size.

The significant interaction of viewpoint change and visibility [ $F(2,38) = 21.41, p < .001, \eta_p^2 = .53$ ] was further examined with Holm-corrected pairwise  $t$  tests ( $p < .05$ ). With visible objects, there was no difference between conditions without and with smooth viewpoint changes, in line with the results of Liu et al. (2005). Performance decreased strongly with abrupt viewpoint changes, replicating the results in Huff et al. (2009). With temporarily invisible objects, smooth viewpoint changes impaired performance, but less so than did abrupt viewpoint changes. The difference between smooth and abrupt viewpoint changes was approximately one target (see Figure 2). When tracking three targets, for example, the participants were able to track two targets with smooth but only one target with abrupt viewpoint changes.

The observed interaction between the viewpoint change and number of targets factors [ $F(4,76) = 3.05, p = .022$ ,



**Figure 1.** Schematic illustration of the viewpoint change (smooth and abrupt) conditions at object invisibility in Experiment 1. The only difference between the two conditions was the kind of viewpoint change of the empty floor plane. Smooth viewpoint changes consisted of 15 frames within 250 msec.

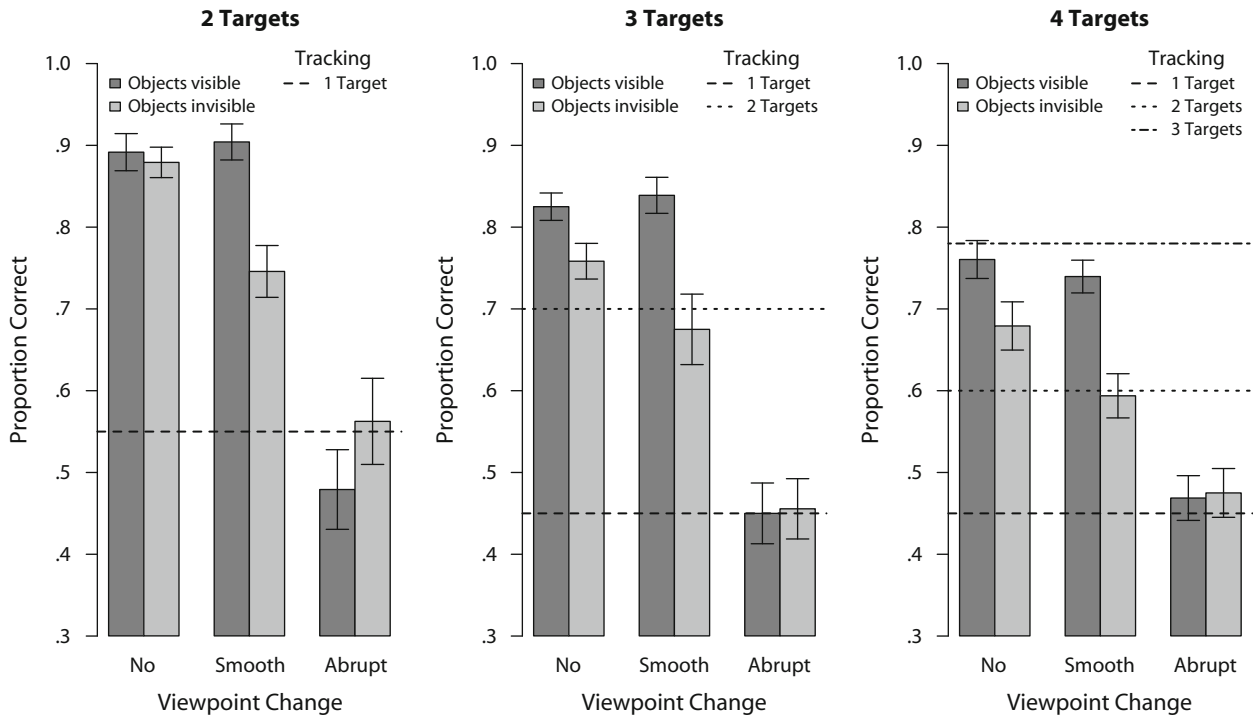


Figure 2. Mean proportions of correct responses (reflecting tracking accuracy) in Experiment 1; error bars indicate standard errors. The horizontal lines depict the tracking performance levels to be expected, according to Hulleman (2005), at tracking capacities of one, two, and three targets.

$\eta_p^2 = .14$ ) was due to the differing performance levels for two, three, and four targets, where performance dropped to similar levels for all numbers of targets with abrupt viewpoint changes. Neither the two-way interaction of invisibility interval and number of targets nor the three-way interaction of viewpoint change, invisibility interval, and number of targets reached significance [ $F(2,38) = 1.89, p = .17, \eta_p^2 = .09$ , and  $F(4,76) < 1$ , respectively]. The main effects of all the variables were significant: Tracking performance decreased with an increasing number of targets [ $F(2,38) = 48.72, p < .001, \eta_p^2 = .72$ ] and with invisibility intervals [ $F(1,19) = 23.75, p < .001, \eta_p^2 = .56$ ]. The main effect of viewpoint change was also significant [ $F(2,38) = 166.28, p < .001, \eta_p^2 = .90$ ].

Better performance with smooth than with abrupt viewpoint changes for temporarily invisible objects supports the hypothesis that continuous information about scene motion can be used to update retained target locations. We suppose that retained target locations were linked to the visible reference frame, which moved continuously during smooth viewpoint changes. Then, with this visible continuous motion, the retained target locations would have been updated—not without any loss in performance, but more successfully than with abrupt scene motion.

In Experiment 1, the floor plane had a checkerboard pattern. As a reference frame, the checkerboard provided visual cues in the vicinity of objects. To test whether these local scene elements are necessary for the updating of retained target locations, we reduced the visual detail of the scene and used a monochromatic wire frame instead

of a checkerboard, as a reference frame in Experiment 2 (see Figure 3).

## EXPERIMENT 2

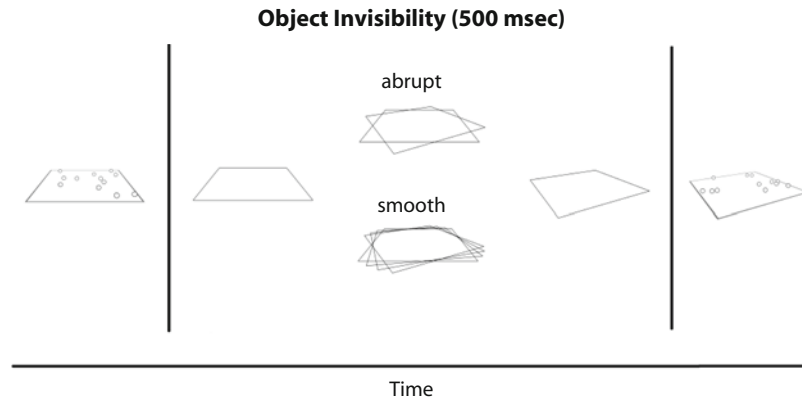
### Method

**Participants.** Twenty students at the University of Tübingen (15 of them female), who were enlisted in the institute’s participant pool but had never before participated in a MOT experiment, participated in the experiment. Their mean age was 23.0 years (20–26 years). All the participants reported normal or corrected-to-normal vision. They received compensation for their participation.

**Apparatus, Stimuli, and Procedure.** All instruments and stimuli and the procedure were exactly the same as those in Experiment 1, except for the floor plane pattern. To reduce scene elements in the vicinity of objects, the light gray and dark gray squares of the checkerboard floor plane were replaced by a single square in the background color, whose borders were indicated by a gray wire frame (see Figure 3). Note that a minimum amount of scene information was necessary to provide continuous information about scene motion in conditions with object invisibility. Otherwise, the abrupt and smooth viewpoint changes would have been indistinguishable.

### Results and Discussion

Mean proportions of correctly identified targets are plotted in Figure 4. As with a checkerboard floor plane, performance was superior with smooth, as compared with abrupt, viewpoint changes even if the objects were invisible during the viewpoint change. Again, the proportion correct scores for all the participants were submitted to a 3 (viewpoint change)  $\times$  2 (object visibility)  $\times$  3 (number of targets) factorial repeated measures ANOVA. The



**Figure 3. Schematic illustration of the viewpoint change (smooth and abrupt) conditions at object invisibility in Experiment 2.**

results replicated the findings of Experiment 1. Again, we observed an interaction between viewpoint change and object invisibility [ $F(2,38) = 15.47, p < .001, \eta_p^2 = .45$ ]. Object invisibility made a clear difference only with smooth viewpoint changes (Holm-corrected pairwise  $t$  tests,  $p < .05$ ). Number of targets induced differing levels of performance without and with smooth viewpoint changes. The levels to which performance dropped differed less with abrupt viewpoint changes, resulting in an interaction between viewpoint change and number of targets [ $F(4,76) = 2.76, p = .034, \eta_p^2 = .13$ ]. Neither the two-way interaction of invisibility interval and number of targets nor the three-way interaction of viewpoint change, invisibility interval, and number of targets reached significance [ $F(2,38) = 1.78, p = .18, \eta_p^2 = .09$ , and  $F(4,76) = 1.76, p = .15, \eta_p^2 = .08$ , respectively]. All three main effects were significant [number of targets,  $F(2,38) = 60.64, p < .001, \eta_p^2 = .76$ ; object invisibility,  $F(1,19) = 9.84, p = .005, \eta_p^2 = .34$ ; viewpoint change,  $F(2,38) = 86.95, p < .001, \eta_p^2 = .82$ ].

The goal of this experiment was to test whether updating the retained locations of invisible targets would be possible with minimized local scene information. The textured floor plane in Experiment 1 may have helped the observers to retain a target's location by relating it to the corresponding square of the checkerboard pattern. With the monochromatic floor plane as a frame of reference in Experiment 2, such a strategy was prevented, and the same pattern of results was still produced. In summary, Experiment 2 showed that the process of spatial updating is not dependent on detailed local scene information in the vicinity of tracked objects. Updating the locations of invisible objects is possible with reduced information about scene motion. However, continuous information about scene motion is critical. Performance was even slightly better than in Experiment 1 with invisible objects, which may have been due to individual differences between participant groups. Or, checkerboard associations may have been used, but the checkerboard pattern may have been disadvantageous because its sudden change at abrupt viewpoint changes distorted retained object locations. In Experiment 3, we

tested this assumption by comparing checkerboard and wire frame floor planes in one experiment.

### EXPERIMENT 3

#### Method

**Participants.** Nineteen students who were recruited from the psychology department of the University of Greifswald (16 of them female) participated in the experiment. Their mean age was 21.8 years (19–29 years). All the participants reported normal or corrected-to-normal vision. They received compensation for their participation.

**Apparatus, Stimuli, and Procedure.** All the instruments and stimuli and the procedure were exactly the same as those in the first two experiments, with the following exceptions. There were always three targets, and in order to improve statistical power, we increased the number of repetitions per condition from 6 to 16. In addition, we introduced the floor plane pattern as an additional within-subjects factor. The factors *floor plane* (checkerboard or wire frame), *viewpoint change* (none, smooth, or abrupt), and *visibility* (always visible or invisible for 500 msec) were all varied within subjects, resulting in 192 experimental trials that were split into two sessions lasting approximately 1 h each. In addition, there were 24 practice trials at the beginning of each session, which were excluded from statistical analysis.

#### Results and Discussion

Mean proportions of correctly identified targets are plotted in Figure 5. As in Experiments 1 and 2, performance was superior with smooth, as compared with abrupt, viewpoint changes even when the objects were invisible during the viewpoint change. Proportion correct scores for all the participants were submitted to a 2 (floor plane)  $\times$  3 (viewpoint change)  $\times$  2 (object visibility) factorial repeated measures ANOVA. Chance performance, defined as the ability to track one target (Hulleman, 2005), was .45. The results replicated the findings of Experiments 1 and 2. Again, we observed an interaction between viewpoint change and object invisibility [ $F(2,36) = 31.56, p < .001, \eta_p^2 = .64$ ]. Across floor planes, object invisibility made a clear difference only with no and smooth viewpoint changes (Holm-corrected pairwise  $t$  tests,  $p < .001$ ). All two- and three-way interactions involving the

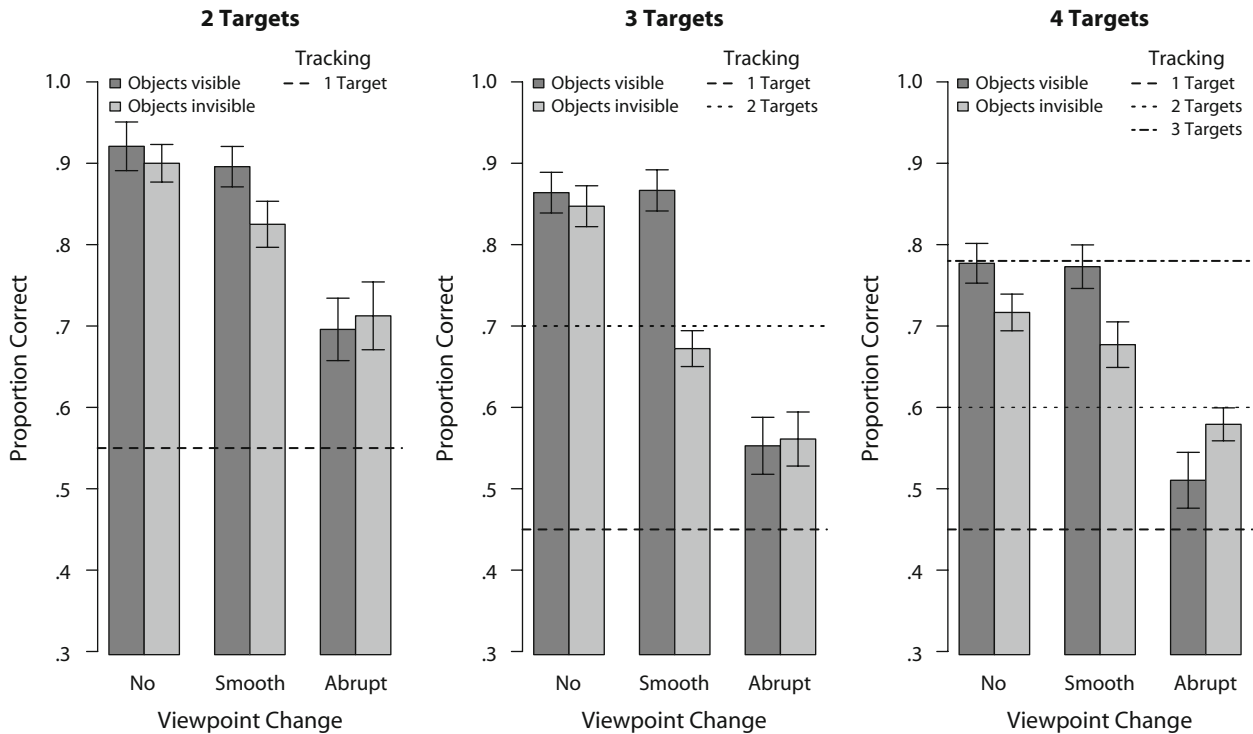


Figure 4. Mean proportions of correct responses (reflecting tracking accuracy) in Experiment 2; error bars indicate standard errors. The horizontal lines depict the tracking performance levels to be expected, according to Hulleman (2005), at tracking capacities of one, two, and three targets.

floor plane factor did not reach significance [ $F_s < 1.79$ ,  $p_s > .18$ ,  $\eta_p^2 < .09$ ]. The main effects of object invisibility [ $F(1,18) = 52.31$ ,  $p < .001$ ,  $\eta_p^2 = .74$ ] and viewpoint change [ $F(2,36) = 180.00$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ] were significant. The main effect of floor plane was not significant [ $F(1,18) < 1$ ].

The differences between the conditions with object invisibility that are apparent in Figure 5 were confirmed in pairwise tests: Performance with smooth viewpoint changes was superior to performance with abrupt viewpoint changes but was impaired in comparison with tracking invisible objects without viewpoint changes (Holm-corrected pairwise  $t$  tests,  $p < .001$ ). This pattern of results reflects the fact that continuous information is important for spatial updating processes and that object displacement during object invisibility impairs visual tracking even if continuous information about scene motion is available.

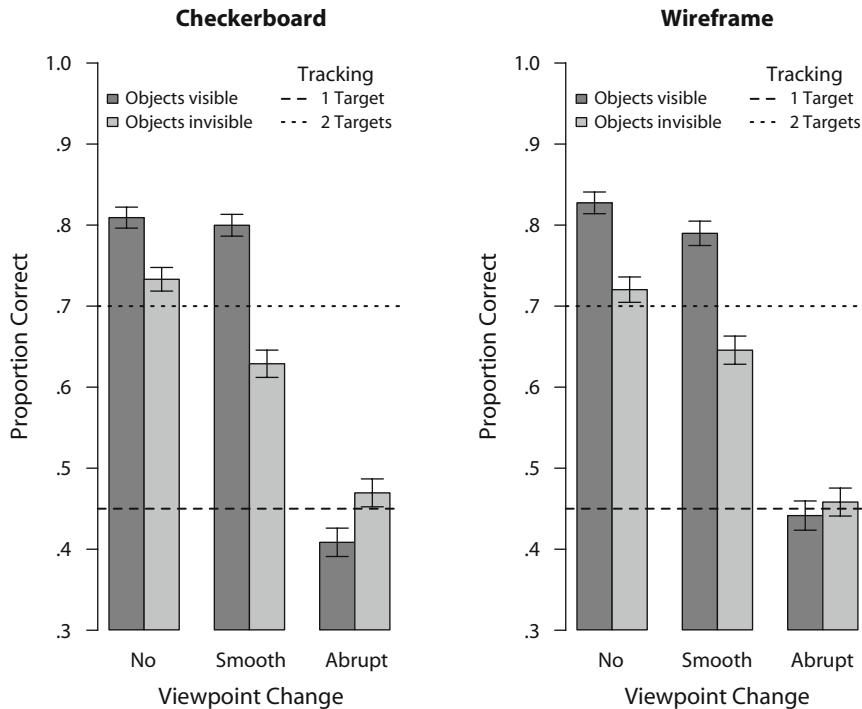
The goal of this experiment was twofold. First, we wanted to directly test whether the difference in tracking performance observed between Experiments 1 and 2 was due to different floor planes. Second, we wanted to replicate the findings of the first two experiments with a more reliable measure. In this experiment, we found no difference between the checkerboard and the wire frame floor plane [ $F(1,18) < 1$ ]. Hence, performance differences between Experiments 1 and 2 were most likely based on individual differences between the participants. Whereas the participants in Experiment 2 were recruited from the insti-

tute’s participant pool and were, therefore, highly familiar with experimental settings in general (excluding MOT), the participants in Experiments 1 and 3 were most likely less familiar with such settings. This might have caused the observed differences between Experiments 1 and 2.

Most important, we were able to replicate findings of the first two experiments with increased statistical power. With 16 instead of 6 replications per condition, tracking performance with smooth viewpoint changes during object invisibility was well above chance level, which was defined as the ability to track one target [ $t_s(18) > 5.41$ ,  $p_s < .001$ ; Hulleman, 2005]. In contrast, visual tracking performance dropped to approximately this level with abrupt changes and temporarily invisible objects.

### GENERAL DISCUSSION

We examined the role of scene information in MOT with smooth and abrupt viewpoint changes during brief intervals of object invisibility. In control conditions, we replicated the previous findings. As compared with the baseline condition with visible objects and no viewpoint change, object invisibility impaired visual tracking in Experiments 1 and 3 (Keane & Pylyshyn, 2006). If the objects were visible throughout the trial, abrupt (Huff et al., 2009), but not smooth, viewpoint changes (Liu et al., 2005) disrupted MOT performance. Not surprisingly, abrupt viewpoint changes disrupted performance as well if the objects were temporarily invisible. The novel evi-



**Figure 5.** Mean proportions of correct responses (reflecting tracking accuracy) in Experiment 3; error bars indicate standard errors. The horizontal lines depict the tracking performance levels to be expected, according to Hulleman (2005), at tracking capacities of one and two targets.

dence obtained in Experiment 1 and corroborated in Experiments 2 and 3 pertains to smooth viewpoint changes that occurred while the objects were invisible. Performance was better than with abrupt viewpoint changes, suggesting that continuous information about scene motion supported updating of the retained locations of invisible objects. Note that those conditions differed only by the kind of floor plane rotation, whereas object behavior was identical.

Our results clearly show that the locations of targets in MOT are retained in relation to scene elements that act as a reference frame (see also Hollingworth, 2007). In the present experiments, the floor plane indicated by a checkerboard pattern or a wire frame provided a salient reference. In common MOT displays, the screen borders and, if presented, a static fixation mark could act as scene reference. Current theories of MOT state no explicit assumptions regarding spatial reference in visual tracking (Cavanagh & Alvarez, 2005; Oksama & Hyönä, 2004; Pylyshyn, 2007). Authors who have studied the tracking of temporarily invisible objects have mentioned two theories. Keane and Pylyshyn (2006) favored the FINST theory (Pylyshyn, 2001, 2007). They supposed that preattentive indices track targets independently of the scene. If an object disappears and reappears, such an index grabs the object next to the location of disappearance. An index can be represented in multiple coordinate systems, such as eye-, head-, and body-centered frames of reference. Active observer motion automatically updates represen-

tations in multiple reference frames (Pylyshyn, 2007). However, interpreting the results of Liu et al. (2005) in the light of the FINST theory requires updating FINSTs within a reference frame even without observer motion; continuous information seems to be sufficient for this process. In addition, our results show that FINSTs of temporarily invisible objects can be updated with respect to a continuously changing reference frame.

The second theory supposed to explain the tracking of invisible objects postulates that the objects' locations right before disappearance including their direction of motion are held offline in memory (Horowitz et al., 2006; see also Fencsik, Horowitz, Klieger, & Wolfe, 2004; Fencsik, Klieger, & Horowitz, 2007). We have presented evidence that such an offline memory must be encoded in scene coordinates. Furthermore, only continuous information about scene motion supports the use of such a memory representation, held offline, for relocating targets. With offline memory, updating of target representations would occur right after a viewpoint change, not concurrently with scene motion.

In the present study, we showed that the spatial representation of temporarily invisible objects was updated along with continuous changes of the reference frame. In principle, this can be accomplished by means of any mechanism linking stored object locations to a reference frame, be it preattentive indices or FINSTs (Pylyshyn, 2001), offline memory (Horowitz et al., 2006), or foci of split attention following targets (Alvarez & Franconeri, 2007).

Our results tie in with those of previous studies on MOT in which the importance of continuous visual information for successful tracking has been demonstrated (Huff et al., 2009; Liu et al., 2005; Seiffert, 2005). In previous studies, abrupt scene motion was confounded with hard onsets of targets at the new locations. With abrupt scene motion, targets were displaced from one frame to another. With smooth scene motion, there were no large target displacements and no hard onsets at new locations. Hard onsets could impair tracking because of attentional capture. The preattentive indices postulated in the FINST theory are also assumed to grab suddenly appearing objects. In the present study, we solved this confound by introducing intervals of object invisibility during both abrupt and smooth viewpoint changes. Thus, there were hard onsets even with smooth viewpoint changes.

Although we observed significant effects of object invisibility in conditions without viewpoint changes in Experiment 3 (Cohen's  $d = 1.05$ ), abrupt viewpoint changes had a greater impact on visual tracking (Cohen's  $d = 1.52$ ; difference between smooth and abrupt changes in conditions with object invisibility). Hence, continuity of scene information is crucial when multiple objects are tracked across viewpoint changes.

Surprisingly, there was a slight edge for conditions with abrupt viewpoint changes and invisible objects, as compared with abrupt viewpoint changes and visible objects. Although this effect was not statistically significant, we observed this difference in each experiment. Due to low performance (near *track one*, which was defined as chance level; Hulleman, 2005), it might be that our experiments were not able to reliably differentiate between those conditions. Further research is needed to clarify whether object invisibility during abrupt viewpoint changes enhances visual-tracking performance. Visual masking (Enns & Di Lollo, 2000) caused by abrupt translations of the visible objects might be a possible explanation for this finding. Another explanation could be that spatial-updating processes (which align the spatial representation after abrupt viewpoint changes) and visual tracking require the same mental resources and cannot be processed in parallel. Object invisibility helped the participants to focus on visual updating and then to switch back to tracking. Further research is needed to clarify whether object invisibility during abrupt viewpoint changes enhances visual-tracking performance.

In the present study, we introduced continuous changes of the environment by smoothly rotating the scenes while the position of the observer remained static. This is the typical scenario when a movie or sports report is watched on TV. As compared with this, tracking multiple objects with moving observers in real or virtual environments adds vestibular movement signals that may interfere with MOT (Thomas & Seiffert, 2009). Acting and navigating in dynamic environments such as road traffic is impossible without a mechanism that adapts retained locations of temporarily invisible objects to continuous changes of the surrounding environment. Such changes can be caused by head and body movements or by motion in a vehicle.

Further research including observer movement is needed to fully understand the complex interplay of object and observer motion. Our study showed that spatial updating is possible for several temporarily invisible objects in parallel and that it benefits from continuous changes of the surrounding scene.

**AUTHOR NOTE**

This research was supported by German Research Foundation (DFG) Grants HU 1510/4-1 and JA 1761/5-1. The first two authors contributed equally to this work. The authors thank Stephan Schwan, Todd Horowitz, Johan Hulleman, Jeremy Wolfe, and an anonymous reviewer for their helpful comments on an earlier version of the manuscript. Correspondence concerning this article should be addressed to M. Huff, Knowledge Media Research Center, Konrad-Adenauer-Str. 40, 72072 Tübingen, Germany (e-mail: m.huff@iwm-kmrc.de).

**REFERENCES**

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, Art. 14), 1-10. doi:10.1167/7.13.14

ALVAREZ, G. A., HOROWITZ, T. S., ARSENIU, H. C., DIMASE, J. S., & WOLFE, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception & Performance*, *31*, 643-667. doi:10.1037/0096-1523.31.4.643a

CAVANAGH, P., & ALVAREZ, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *9*, 349-354. doi:10.1016/j.tics.2005.05.009

ENNS, J. T., & DI LOLLO, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*, 345-352. doi:10.1016/S1364-6613(00)01520-5

FENCSEK, D. E., HOROWITZ, T. S., KLIEGER, S. B., & WOLFE, J. M. (2004). Target reacquisition strategies in multiple object tracking [Abstract]. *Journal of Vision*, *4*(8), 370a. doi:10.1167/4.8.370

FENCSEK, D. E., KLIEGER, S. B., & HOROWITZ, T. S. (2007). The role of location and motion information in the tracking and recovery of moving objects. *Perception & Psychophysics*, *69*, 567-577.

FLOMBAUM, J. I., SCHOLL, B. J., & PYLYSHYN, Z. W. (2008). Attentional resources in visual tracking through occlusion: The high-beams effect. *Cognition*, *107*, 904-931. doi:10.1016/j.cognition.2007.12.015

HOLLINGWORTH, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception & Performance*, *33*, 31-47. doi:10.1037/0096-1523.33.1.31

HOROWITZ, T. S., BIRNKRANT, R. S., FENCSEK, D. E., TRAN, L., & WOLFE, J. M. (2006). How do we track invisible objects? *Psychonomic Bulletin & Review*, *13*, 516-523.

HUFF, M., JAHN, G., & SCHWAN, S. (2009). Tracking multiple objects across abrupt viewpoint changes. *Visual Cognition*, *17*, 297-306. doi:10.1080/13506280802061838

HULLEMAN, J. (2005). The mathematics of multiple object tracking: From proportions correct to number of objects tracked. *Vision Research*, *45*, 2298-2309. doi:10.1016/j.visres.2005.02.016

JEFFERIES, L. N., WRIGHT, R. D., & DI LOLLO, V. (2005). Inhibition of return to an occluded object depends on expectation. *Journal of Experimental Psychology: Human Perception & Performance*, *31*, 1224-1233. doi:10.1037/0096-1523.31.6.1224

KEANE, B. P., & PYLYSHYN, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, nonpredictive function. *Cognitive Psychology*, *52*, 346-368. doi:10.1016/j.cogpsych.2005.12.001

KLEIN, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*, 138-147. doi:10.1016/S1364-6613(00)01452-2

LIU, G., AUSTEN, E. L., BOOTH, K. S., FISHER, B. D., ARGUE, R., REMPEL, M. I., & ENNS, J. T. (2005). Multiple-object tracking is based on scene, not retinal, coordinates. *Journal of Experimental Psychology: Human Perception & Performance*, *31*, 235-247. doi:10.1037/0096-1523.31.2.235



- 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**, 631-671. doi:10.1080/13506280344000473
- POSNER, M. I., & COHEN, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- PYLYSHYN, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, **80**, 127-158. doi:10.1016/S0010-0277(00)00156-6
- PYLYSHYN, Z. W. (2007). *Things and places: How the mind connects with the world*. Cambridge, MA: MIT Press.
- PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179-197. doi:10.1163/156856888X00122
- SCHOLL, B. J., & PYLYSHYN, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, **38**, 259-290. doi:10.1006/cogp.1998.0698
- SEIFFERT, A. E. (2005). Attentional tracking across display translations [Abstract]. *Journal of Vision*, **5**(8), 643a. doi:10.1167/5.8.643
- THOMAS, L., & SEIFFERT, A. (2009). Self-motion influences multiple-object tracking in a virtual environment [Abstract]. *Journal of Vision*, **9**(8), 242a. doi:10.1167/9.8.242
- TIPPER, S. P., DRIVER, J., & WEAVER, B. (1991). Short report: Object centered inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, **43A**, 289-298.
- YI, D.-J., KIM, M.-S., & CHUN, M. M. (2003). Inhibition of return to occluded objects. *Perception & Psychophysics*, **65**, 1222-1230.

(Manuscript received August 6, 2009;  
revision accepted for publication October 18, 2009.)