Does action disrupt
Multiple Object Tracking (MOT)?

Ian M. Thornton[1] & Todd S. Horowitz[2]

[1]*Department of Cognitive Science, Faculty of Media and Knowledge Sciences,
University of Malta, Malta.*
[2]*National Cancer Institute, National Institutes of Health,
Bethesda, MD, USA*

While the relationship between action and *focused* attention has been well-studied, less is known about the ability to *divide* attention while acting. In the current paper we explore this issue using the multiple object tracking (MOT) paradigm (Pylyshyn & Storm, 1988). We asked whether planning and executing a display-relevant action during tracking would substantially affect the ability track and later identify targets. In all trials the primary task was to track 4 targets among a set of 8 identical objects. Several times during each trial, one object, selected at random, briefly changed colour. In the baseline MOT trials, these changes were ignored. During active trials, each changed object had to be quickly touched. On a given trial, changed objects were either from the tracking set or were selected at random from all 8 objects. Although there was a small dual-task cost, the need to act did not substantially impair tracking under either touch condition.

Keywords: Divided Attention, Action, Multiple Object Tracking, MOT, Tracking, Mobile Devices

Multiple object tracking (MOT, Pylyshyn & Storm, 1988) has become a standard paradigm for examining the ability to divide attention in dynamic environments (for a review see, Scholl, 2009). In a typical display, observers are shown a set of identical objects, half identified as targets (usually by briefly highlighting or blinking them) and half as distractors. The display is set in motion and all of the (now identical) objects follow independent, random trajectories. At the end of the tracking period, the motion stops and the observer is asked to identify the targets. The dependent measure is usually the inferred proportion of targets correctly tracked (Hulleman, 2005). While the task appears

quite demanding, most participants are able to successfully track 3–5 items. Beyond 3–5 items, however, tracking ability appears to be severely limited in most experimental situations (cf. Franconeri, Jonathan, & Scimeca, 2010). Several explanations have been proposed for this limit, including a fixed set of virtual pointers (Pylyshyn, 1989, 2009), flexible attentional resources (Alvarez & Franconeri, 2007), and limitations in working memory capacity (Allen, Mcgeorge, Pearson, & Milne, 2006).

In the current paper, we ask if concurrently performing an action during tracking affects the ability to identify items at the end of a trial. Although action has been examined in great detail within the context of focused attention (Allport, 1987; Hommel, 2010; Humphreys et al., 2010; Riddoch, Humphreys, Edwards, Baker, & Willson, 2003; Schneider & Deubel, 2002; Symes, Tucker, Ellis, Vainio, & Ottoboni, 2008), to our knowledge, much less is known about the consequences of acting while attention is divided. While there are several real-world scenarios where successful multi-tasking performance might suggest little impact of action on divided attention — for example in the context of team sports or complex control scenarios, such as CCTV or air traffic control centres — there are other reasons to suspect that competition for limited resources could significantly modulate performance.

For example, both the "selection for action" hypothesis (Allport, 1987) and the "pre-motor theory of attention" (Rizzolatti, Riggio, Dascola, & Umiltá, 1987; Rizzolatti, Riggio, & Sheliga, 1994) predict that motor preparation and execution result in a mandatory reallocation of attention. Consistent with these ideas, there is now considerable evidence that executing an action, either with the eyes (Born, Mottet, & Kerzel, 2014; Deubel & Schneider, 1996; Hunt & Kingstone, 2003; Kowler, Anderson, Dosher, & Blaser, 1995; Shepherd, Findlay, & Hockey, 1986) or hands (Bekkering & Neggers, 2002; Deubel, Schneider, & Paprotta, 1998; Eimer, Van Velzen, Gherri, & Press, 2006; Schiegg, Deubel, & Schneider, 2003), causes a shift of attention towards the target object/location. In the context of MOT, requiring the participant to touch an individual object in the display may thus result in a significant and unpredictable (from the perspective of on-going tracking) redistribution of attention. The purpose of the current study was to examine whether such a redistribution is enough to break, or substantially reduce the ability to track multiple objects.

Recently, we introduced a new task designed to examine the ability to *control* multiple objects, rather than to simply track them for identity (Thornton, Bülthoff, Horowitz, Rynning, & Lee, 2014). This *interactive* Multiple Object Tracking (iMOT) task was directly inspired by mobile app games, such as such as Flight Control (Firemint Pty Ltd) and Harbor Master (Imangi Studios, LLC), and was itself implemented on an iPad. Participants used touch control to guide objects, and the goal of the task was to avoid collisions during a fixed time period. We found that with display parameters in which participants could

track 4 out of 8 items in an MOT task, they could successfully control 6 items without collision during iMOT. Performance on standard MOT and on iMOT was strongly positively correlated (r = .72) in the same individuals (Thornton et al., 2014).

Of most relevance to the current topic, in one experiment participants were required to perform both MOT and iMOT at the same time. Our prediction was that MOT performance would break down as planning and executing the iMOT control movements would reduce available attentional resources for tracking (Allport, 1987; Rizzolatti et al., 1994). Contrary to this prediction, MOT performance actually improved slightly under dual-task conditions (Exp3; Thornton et al., 2014). On first pass, this would seem to suggest some level of independence between focused and divided attentional resources, and may even hint at some form of "synergistic lock". That is, focusing attention on and/or interacting with one item in a tracked set might temporarily improve the ability to localise all of the other members of a set being tracked with divided attention.

There were, however, two design issues with this original experiment that suggest caution in generalising to other MOT scenarios. To begin with, we did not control for strategic reallocation of resources between the two tasks, and indeed, iMOT collisions increased under dual-task conditions. Participants were able to choose the moment to act, and may simply have waited until a collision was impending with one of the tracked objects, requiring little if any reallocation of attention. Perhaps more importantly, only objects that were in the tracking set responded to touch control. This was done to avoid participants "herding" target and distractor items to different sides of the screen, but it also clearly provided a means where "lost" targets could be recovered.

The current experiment was designed to address these issues and to provide a clear and simple test of the impact of acting while performing MOT. As illustrated in Figure 1, on all trials the primary task was to track 4 targets among a set of 8 objects. Five times during each trial, one object, selected at random, changed colour for 2 seconds. In a baseline MOT block of 10 trials, these changes were ignored. During the dual-task "action" block of 20 trials, each changed object had to be quickly touched. To ensure that participants immediately allocated attention to this secondary task, the trial was aborted if the colour singleton was not touched within the two second period. Except for returning to its original colour, a touched object did not change its behaviour in any way following the intervention.
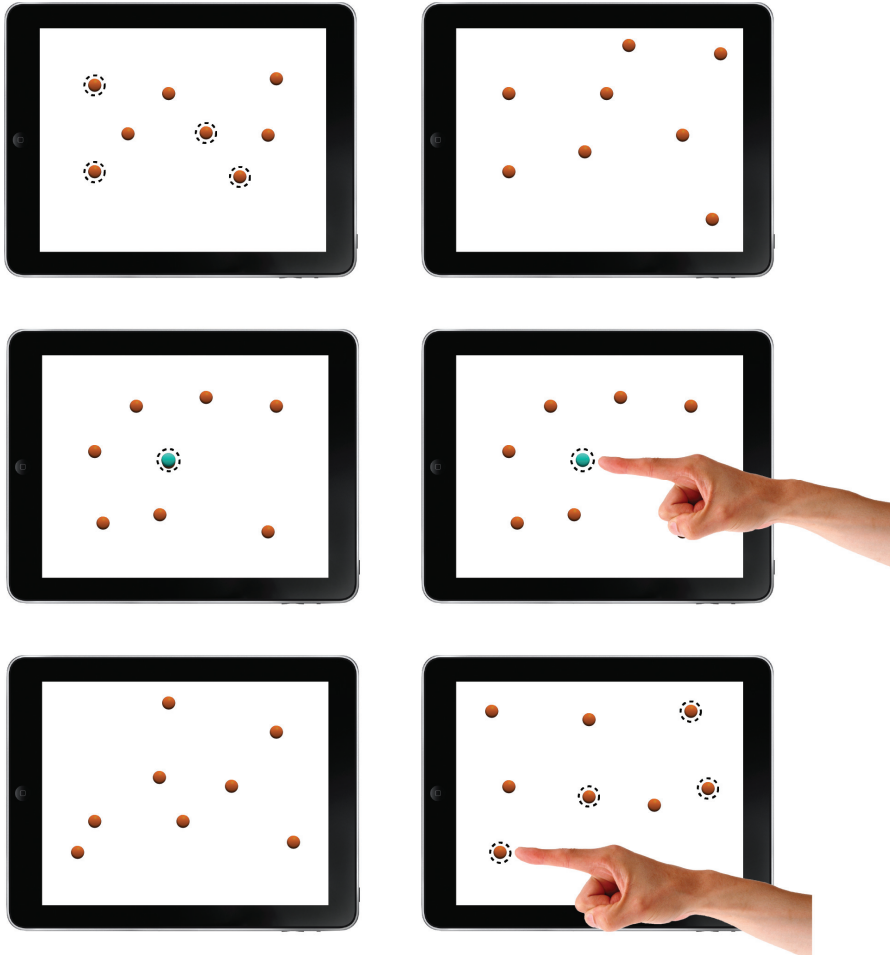
Figure 1. Schematic timeline of a typical trial. Top row: After a 1 second preview, 4 objects began blinking for 2 seconds, identifying them as targets (broken circle and white background for illustration only). All of the (now identical) objects then began to move on random trajectories. Middle row: Five times during every trial one object would change colour for 2 seconds. In baseline MOT trials these changes were ignored. In action trials, the object had to be rapidly touched, after which the colour change was cancelled and tracking continued. Bottom row: At the end of the tracking period participants identified the targets by touching them. Each touched item changed colour. See text for more details.

For 10 of the dual-task trials, the singleton object was selected randomly from the tracking set and for the remaining 10 trials, it was randomly selected from all 8 objects. These two types of trial were randomly intermixed and participants had no independent cues as to the type of trial. This manipulation was made so that on half of the trials, at least some touches would require focused attention to be allocated away from the tracking set.

Method

Participants

A total of 12 participants (9 females, mean age 22.2 years, SD = 6.4) from the University of Malta community took part in this experiment in return for payment of €4. The sample size was determined prior to data collection based on values typical to the field and to previous studies in our laboratories (e.g., Thornton et al., 2014). All participants had normal or corrected to normal vision, and were naïve as to the purpose of the research until data collection was complete. All participants gave written informed consent, and all aspects of the procedure were reviewed and approved by the Research Ethics Committee of the Faculty of Media & Knowledge Sciences, University of Malta, conforming to the principles of the Declaration of Helsinki.

Equipment

The experiment was conducted using a first generation iPad with a screen dimension of 20 x 15 cm and a resolution of 1024 x 768 pixels. Participants were instructed to cradle the iPad (in landscape orientation) in their left arm, with the fingers of their left hand grasping the furthest edge of the device. They were required to respond to objects using the index finger of their right hand. While viewing distance was not fixed, we estimated that it was approximately 50 cm from screen surface to eyes. The experiment was run in a quiet environment under low lighting conditions with no overhead lights, in order to minimize screen glare.

Stimuli

Objects were identical orange spheres with a diameter of 52 pixels (1.2°). The objects were shaded so that they appeared to be lit from above to provide an impression of 3D and to help segment them from the uniform black background. At the start of each trial, the objects were stationary and were distributed randomly across the display. Object position was determined by sparsely populating a 5 x 5 invisible grid, and then perturbing each object a random distance from the centre of the cell. This was done to ensure that objects did not initially overlap. After a 1 second preview, 4 of the objects began to blink, identifying them as targets. After 2 seconds, the blinking stopped and all 8 (now identical) objects began to move in random directions sampled from the full 360° in 1° increments. Objects always moved at a constant speed of approximately 2°/s, changing direction after a variable path length of between 200 pixels (3.9°) and 300 pixels (5.9°). Path length and direction was constrained so that objects remained within the display area. If two objects collided they passed through each other and did not bounce.

In all three conditions, 5 colour change events occurred during each trial, with a single object changing from orange to blue. During baseline MOT trials these events always lasted 2 seconds, and then the object returned to its original colour. During the dual-task action block, the colour returned to original as soon as the object was touched. Failure to touch the object during the 2 second period resulted in the trial being aborted and replaced later in the block. For all types of trial, the next colour change event was scheduled at the end of a 2 second period, timed from colour-change onset, and occurred at a variable gap of between 2–3 seconds.

Task

The primary task in all conditions was to track the 4 target objects, as in standard MOT. In all types of trial the objects were in motion for approximately 30 seconds, at which point the participant was required to identify the 4 target objects by touching them. These responses

were self-paced – there was no additional cue to respond – and we recorded the time of each subsequent touch relative to the end of object motion. Touched targets changed colour from orange to purple. During training trials, feedback was given on tracking performance by briefly blinking the correct 4 target items. No feedback was given during the experimental trials. At the end of the trial the display faded and a blank, self-paced pause screen was entered. Participants were able to initiate the next trial by clicking on a "Continue" button.

There were two dual-task conditions, "touch-targets" and "touch-all". In both conditions an immediate touch response was required to an object colour change. If the trial was aborted because a touch did not occur within 2 seconds, the display immediately faded to black and the pause screen appeared. The omission of the MOT response phase clearly marked aborted trials, and no other explicit feedback on touch errors was provided.

In the touch-targets condition these 5 colour-change events were always drawn from the tracking set. In the touch-all condition, the changed object was drawn from the tracking or distractor set with equal probability within a single trial. Note that during the baseline block, when no touch response was required, colour-change events were sampled according to the touch-all schedule.

Procedure

Participants were run in individual sessions. Each session began with a brief familiarization phase, where the iPad and the basic display and control components of the task were explained. Participants always began with the 10 MOT baseline trials, and so they were initially told to ignore the colour-change events. Instruction and practice thus focused on the MOT components of the task, and participants typically completed 2 or 3 demo trials during this phase. The familiarization phase and the MOT baseline trials typically took less than 10 minutes to complete.

When the MOT baseline trials had been completed, the dual-task touch aspects of the experiment were explained. Further practice trials, typically 2 or 3, were also given with this new component. The 20 trials of the dual-task action block were then completed, with this phase typically taking a further 10 minutes. The two types of touch trial – 10 touch-target and 10 touch-all — were randomly interleaved in the design with no explicit cues provided that would allow them to be distinguished.

Analysis

The main dependent measure was the number of correctly tracked objects. Data were analysed using one-way analysis of variance (ANOVA), with touch condition (baseline, touch-targets, touch-all) as the repeated measure. Pair-wise planned comparisons were used to explore differences between each condition. We also examined two response time measures. The first was the speed with which MOT targets were identified at the end of the trial. This was a cumulative measure with each trial yielding 4 response times, one for each target identification. These responses were analysed using a 4 (Target) x 3 (Touch Condition) repeated measures ANOVA. The second measure was restricted to the two touch conditions and examined the average response time to the colour singleton objects. A paired t-test was used to compare responses in these two conditions.

We note that across all participants there were only 7 instances of trials being aborted due to failure to respond to the colour singletons within 2 seconds. As these errors were so few, and as aborted trials were replaced in the design, we will not discuss them further. Where appropriate Greenhouse-Geisser corrections were applied during repeated measures ANOVA to adjust for violations of the sphericity assumption.

Results

The main results from this experiment are summarized in Figure 2. In all conditions participants were able to successfully identify at least 3 target objects, a level of performance that is comparable with standard MOT tracking. Although the ability to track appears not to have been severely disrupted by the addition of a second, action-related task, there was a small but consistent decrement in performance relative to the MOT-baseline, giving rise to a significant main effect of condition, $F(2,22) = 31.7$, MSE = 0.04, p <0.001, $\eta^2 = 0.74$. As illustrated in Figure 2, pairwise comparisons of the baseline condition (M = 3.77, SE = 0.08) to both the Touch-target (M = 3.36, SE = 0.12) and Touch-all (M = 3.12, SE = 0.12) conditions were significant at p <.001 level, whereas comparisons of the two action conditions to each other were significant at the p <.05 level.
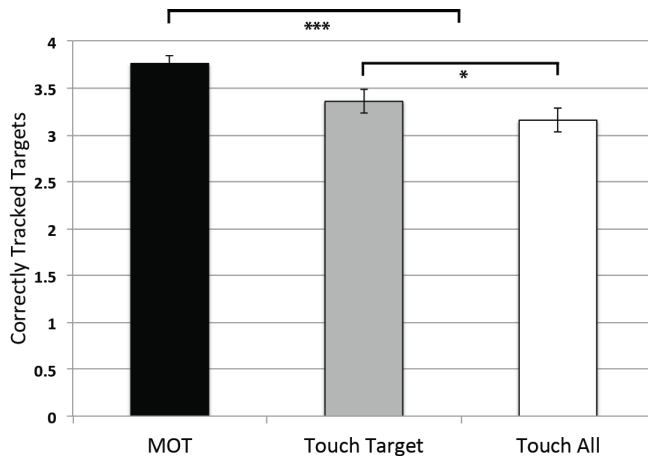

Figure 2. Mean tracking accuracy in each condition. Error bars indicate 1 standard error of the mean.

Figure 3 summarises the data for the MOT response times. There was a main effect of Target, $F(3,33) = 360.20$, MSE = 0.05, p <0.001, $\eta2 = 0.97$, reflecting the sequential nature of the responses. There was also a main effect of Touch Condition, $F(2,22) = 10.34$, MSE = 0.16, p <0.01, $\eta2 = 0.46$. As can be seen in Figure 3, this effect reflects overall slower responses in the Baseline Condition (M = 1622 ms, SE = 100), compared to both the Touch-target (M = 1330 ms, SE = 143) and Touch-all (M = 1271 ms, SE = 117) conditions. Pairwise comparison confirmed these differences were significant at the p <.05 level, but there was no difference between the two touch conditions. The Target x Touch Condition interaction was not significant, $F(6,66) = 2.73$, MSE = 0.04, n.s., $\eta2 = 0.19$.
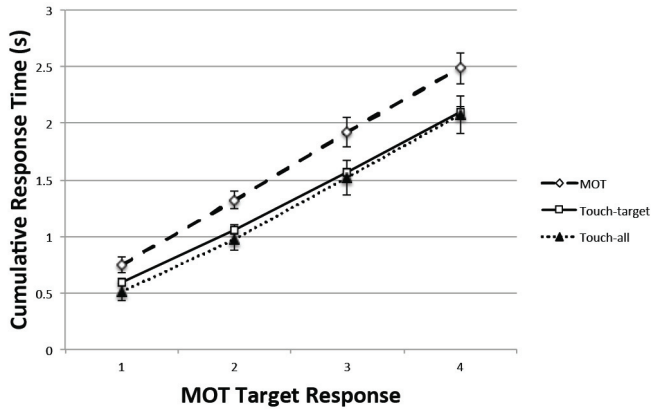
Figure 3. Mean response time for identifying MOT targets at the end of tracking as a function of condition. Error bars indicate 1 standard error of the mean.

Finally, a comparison of the average time to touch the colour singleton during tracking in the Touch-target (M = 822 ms, SE = 27) and Touch-all (M = 830 ms, SE = 26) conditions, indicated that they did not significantly differ from each other, t(11) = 1.0, n.s.

Discussion

The current experiment set out to explore whether performing simple actions during an MOT trial would substantially disrupt the ability to identify target objects. As the planning and execution of movement is known to involve a shift of focused attention towards action targets, our question was essentially whether it is possible to both divide and focus attention at the same time. The results seem quite clear. Although there was a measurable reduction in tracking performance during dual-task action trials, at all times participants were able to track at least 3 out of 4 target objects.

This result is largely in agreement with our previous finding (Thornton et al., 2014) where participants were able to control items at the same time as tracking them. The major difference between these results and our previous findings is that we found a small decrement in MOT performance under dual-task conditions, rather than a small improvement. This suggests that additional, control-related cues may well have helped participants identify targets in our previous task. However, another possibility is that the currant dual-task results reflect a speed-accuracy trade-off, even though the MOT response comes at the end of the trial and is not speeded. That is, RTs when identifying targets were consistently faster in dual-task than in baseline trials, raising the possibility that participants sacrificed reporting accuracy – as opposed to tracking accuracy — in order to respond more quickly (thus reducing total time spent doing the experiment). This would mean that we were overestimating the effect of the dual-task on accuracy.

Conversely, as the dual-task block always followed the baseline block, the RT speed-up could also reflect simple practice effects. We chose a fixed order for the two blocks of trials in order to ensure that the primary MOT task was well-established before trying to disrupt it. Practice effects could have inflated the dual-task performance relative to the initial baseline, meaning that we might be underestimating the dual-task cost. Overall, however, there can be little doubt that any dual-task decrement is small in absolute terms. Taken together then, the current results, together with the findings of Thornton et al., (2014), allow us to claim with some certainty that action does not appear to substantially affect MOT performance under the examined conditions.

The selection-for-action hypothesis suggests that planning and executing an action, such as a pointing movement to a moving object, leads to a shift of attention to the target of the movement. Since MOT is an attention-demanding task, disrupting the distribution of attention should seriously impair tracking, yet we observed only modest reductions in performance. How can we explain this?

We did not directly measure whether attention was actually deployed to the probe object, so it is possible that no shift occurred in this experiment. Studies of the selection-for-action hypothesis (e.g., Bekkering & Neggers, 2002; Eimer et al., 2006) typically measure the spontaneous deployment of attention in the absence of a competing attentional task. Here, the priority assigned to the tracking task may have pre-empted or countermanded the action-driven shift.

A second possibility is that the response probe did engender an attentional shift, but since participants were dividing their attention among multiple objects, the probe attracted only a proportional share of attention, rather than all of the participant's available resources. We can understand this scenario easily if we think in terms of Pylyshyn's (2001, 2007) concept of FINSTs, or visual indexes (see also Alvarez & Franconeri, 2007 's FLEX model). In MOT, these indexes act like figurative fingers that constantly point to the targets. If I have five indexes, it is a simple matter to shift one of the indexes from tracking to the response probe, while still having four indexes to track the four targets. Alternatively, we can think of attention as a continuous resource, like energy or money (Horowitz & Cohen, 2010). In this case, before the probe shows up, each target gets 25% of the available resource. When the action system calls for a shift of attention, 5% is peeled off from each target so that 20% can be directed to the probe. This would lead to slightly less precise information available about target position and motion, thus explaining the modest decrement in performance that we observed.

We must also consider the possibility that the MOT task did not fully occupy the participants' attentional resources. We asked participants to track four targets because this is near the capacity of the typical MOT participant. We expected that participants would be near, but not over the limit of their abilities here. However, tracking capacity varies among individuals, and is dependent on speed as well as load (Alvarez & Franconeri, 2007), so we may have underestimated our participants' capabilities. Indeed, performance in the baseline condition was quite good, so participants may have had spare attentional capacity. Note that this is not mutually exclusive with the divided

attention hypothesis described above. In fact, in order for the presence of spare attentional capacity to be relevant, it has to be the case that participants can divide attention between MOT and action, and that pointing to the probe could not demand a complete redistribution of attention. Thus, this is a subset of the divided attention hypothesis.

A third option is that the requirement to act may in fact have led to a complete shift of attention to the probe, but that attention was then returned to the MOT targets after a brief interval. Consistent with this idea, Deubel & Schneider (2003) have shown that attention may be quickly withdrawn from action targets during the planning/execution of hand movements, although the same is not true for eye movements. Similarly, there is some evidence that participants can successfully track multiple targets even when attention is briefly withdrawn from the task for a fraction of a second, which may be accomplished either by remembering the positions of the targets before the shift (Keane & Pylyshyn, 2006) or by predicting where the targets will be at the end of the interval (Fencsik, Klieger, & Horowitz, 2007).

Further research could help us disambiguate these possibilities. Psychophysical measures of attention to the probe item and studies of the distribution of eye fixations around the time of the probe and the pointing movement would be informative. Given that both tracking (Drew, Horowitz, Wolfe, & Vogel, 2011; Drew & Vogel, 2008) and attentional shifts (Eimer, 1996; Luck & Hillyard, 1994) have distinct event-related potential signatures, electrophysiological studies may be able to shed some light on the attentional dynamics here. It would also be informative to use adaptive methods to adjust the difficulty of tracking below and above each individual's threshold in this task. This would shed some light on the question of whether action can divert attention already devoted to an ongoing task, or merely pulls "spare" attentional resources. More interestingly, such a design might reveal that participants with different attentional capacities adopt different strategies when faced with the conflict between passive tracking and action.

Conclusions

In our everyday life, we are continually scanning and interacting with the world around us. In the laboratory, however, these activities are generally compartmentalized. When we study attention, we minimize the action component, usually saving responses for the end of the trial, using them to retroactively infer the dynamics of attention. When we study action, we usually do not provide a concurrent attention task. Yet clearly our brains must have evolved to integrate these activities.

We opened the paper with the question of whether action disrupts ongoing attentional tracking. The answer so far is a qualified "no". We observed minimal reductions in tracking performance when participants were asked to respond to intermittent probes by touching the display. Further work will be necessary

to determine whether this finding holds when participants are pushed to the limit, when tracking or action are more demanding. However, we feel that this experiment provides a convincing illustration that attention and action can be efficiently synthesized.

References

Allen, R., Mcgeorge, P., Pearson, D. G., & Milne, A. (2006). Multiple-target tracking: A role for working memory? *The Quarterly Journal of Experimental Psychology*, *59*(6), 1101–1116. http://doi.org/10.1080/02724980543000097

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Erlbaum.

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

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, *13*(4), 370–374. http://doi.org/10.1111/j.0956–7976.2002.00466.x

Born, S., Mottet, I., & Kerzel, D. (2014). Presaccadic perceptual facilitation effects depend on saccade execution: Evidence from the stop-signal paradigm. *Journal of Vision*, *14*(3), 7–7. http://doi.org/10.1167/14.3.7

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.

Deubel, H., & Schneider, W. X. (2003). Delayed saccades, but not delayed manual aiming movements, require visual attention shifts. *Annals of the New York Academy of Sciences*, *1004*(1), 289–296. http://doi.org/10.1196/annals.1303.026

Deubel, H., Schneider, W. X., & Paprotta, I. (1998). Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition*, *5*(1–2), 81–107.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *The Journal of Neuroscience*, *31*(2), 659–668. http://doi.org/10.1523/JNEUROSCI.1339–10.2011

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *The Journal of Neuroscience*, *28*(16), 4183–4191. http://doi.org/10.1523/JNEUROSCI.0556–08.2008

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234. http://doi.org/10.1016/0013–4694(96)95711–9

Eimer, M., Van Velzen, J., Gherri, E., & Press, C. (2006). Manual response preparation and saccade programming are linked to attention shifts: ERP evidence for covert attentional orienting and spatially specific modulations of visual processing. *Brain Research*, *1105*(1), 7–19. http://doi.org/10.1016/j.brainres.2005.10.060

Fencsik, 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*(4), 567–577. http://doi.org/10.3758/BF03193914

Franconeri, S. L., Jonathan, S., & Scimeca, J. (2010). Tracking multiple objects is limited only by object spacing, not speed, time, or capacity. *Psychological Science*, *21*(7), 920–925. http://doi.org/10.1177/0956797610373935

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. In B. J. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 121–140). Cambridge, MA: MIT Press.

Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking is limited by a graded resource. *Attention, Perception & Psychophysics*, *72*(7), 1765–1775. http://doi.org/10.3758/APP.72.7.1765

Hulleman, J. (2005). The mathematics of multiple object tracking: From proportions correct to number of objects tracked. *Vision Research*, *45*(17), 2298–2309. http://doi.org/10.1016/j.visres.2005.02.016

Humphreys, G. W., Yoon, E. Y., Kumar, S., Lestou, V., Kitadono, K., Roberts, K. L., & Riddoch, M. J. (2010). The interaction of attention and action: From seeing action to acting on perception. *British Journal of Psychology*, *101*(2), 185–206. http://doi.org/10.1348/000712609X458927

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: linked or independent? *Cognitive Brain Research*, *18*(1), 102–105. http://doi.org/10.1016/j.cogbrainres.2003.08.006

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*(4), 346–368. http://doi.org/10.1016/j.cogpsych.2005.12.001

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000–1014. http://doi.org/10.1037/0096–1523.20.5.1000

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

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1–2), 127–158.

Pylyshyn, Z. W. (2007). *Things and Places: How the Mind Connects with the World*. MIT Press.

Pylyshyn, Z. W. (2009). The empirical case for bare demonstratives in vision. In R. Stainton & C. Viger (Eds.), *Compositionality, context and semantic values* (pp. 255–274). Dordrecht: Springer.

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

Riddoch, M. J., Humphreys, G. W., Edwards, S., Baker, T., & Willson, K. (2003). Actions glue objects but associations glue words: Neuropsychological evidence for multiple object selection. *Nature Neuroscience*, *6*(1), 82–89. http://doi.org/10.1038/nn984

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1), 31–40.

Rizzolatti, G., Riggio, L. and Sheliga, B.M. (1994). Space and selective attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 231-265). Cambridge, MA: MIT Press.

Schiegg, A., Deubel, H., & Schneider, W. (2003). Attentional selection during preparation of prehension movements. *Visual Cognition*, *10*(4), 409–431. http://doi.org/10.1080/13506280244000140

Schneider, W. X., & Deubel, H. (2002). Selection-for-perception and selection-for-spatial-motor action are coupled by visual attention: A review of recent findings and new evidence from stimulus-driven saccade control. In W. Prinz & B. Hommel (Eds.), *Attention and*

performance XIX: Common mechanisms in perception and action (pp. 609–627). Oxford: Oxford University Press.

Scholl, B. J. (2009). What have we learned about attention from multiple object tracking (and vice versa) ? In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and pylyshyn* (pp. 49–78). Cambridge, MA: MIT Press.

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, *38*(3), 475–491.

Symes, E., Tucker, N., Ellis, R., Vainio, L., & Ottoboni, G. (2008). Grasp preparation improves change detection for congruent objects. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 854–871. http://doi.org/10.1037/0096-1523.34.4.854

Thornton, I. M., Bülthoff, H. H., Horowitz, T. S., Rynning, A., & Lee, S.-W. (2014). Interactive Multiple Object Tracking (iMOT). *PloS One*, *9*(2), e86974. http://doi.org/10.1371/journal.pone.0086974