Available at http://ruccs.rutgers.edu/pylyshyn/VSS-poster-2012.pdf

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Effect of Occlusion and Landmarks on Single Object Tracking During Disrupted Viewing

Meriam Naqvi[1,2], Kevin Zish[1], Ronald Planer[1,3], Deborah Aks[1], Zenon W. Pylyshyn[1]

Rutgers Center for Cognitive Science[1], Rutgers Department of Biological Sciences[2], Rutgers Department of Philosophy[3]

Abstract

A great deal has been learned in the past several decades about how we track multiple independently-moving objects (MOT), including the conditions under which one can track objects that disappear briefly (i.e., during disrupted viewing). What has been lacking is a detailed study of what happens when objects go behind an occluding surface (with a few exceptions: Pylyshyn & Cohen, 1999; Flombaum, et al., 2008). Here we begin a study of the task of tracking a single object that moves behind an occluding surface. Observers were asked to tap on a screen when a sound cue was heard to indicate where they believe the moving object was at the time of the sound cue. We measured the accuracy of this response when the moving object was occluded vs. when visible, and when there were landmarks along the route. We also recorded eye movements to investigate whether gaze may play a role in imagined object-tracking. Method: Thirty-five participants tracked a square moving from left-to-right on a display screen, and twenty-five tracked the square moving right-to-left (during 20 five-second trials for 2 unmarked and 2 landmark conditions consisting of a row of asterisks along the extrapolated trajectory). Subjects selected the position of the tracked (but hidden) object with their finger when signaled by a randomly presented sound probe. Results: Most accurate localization occurs when the object is always visible. In contrast to the finding reported in Pylyshyn & Cohen, 1999, (for significantly different conditions - performed in total darkness), there was only a negligible effect of landmarks. When the object was occluded, gaze and touch localization undershoot actual target position regardless of movement direction or landmark presence. We found that response-lag is greater for gaze, except when the probe onset occurred early in the motion path (<1.9 s). We will discuss how lag-bias may reflect coding of object position in eye-movement system and guide imagined localization and tracking accuracy. Furthermore, we will describe an unexpected finding, suggesting that our eyes may serve as a "place-holder" to maintain the position of tracked non-visible objects (i.e., Immediately following sound probe, eyes tend to stay in approximately same position until response is made).

Methods

- •60 participants tracked a single red-square (1.1°) that moved horizontally 35° across a (38°) black computer screen.
- •Square either moved left-to-right (LR), or right-to-left (RL) across the screen.
- •4 second tracking period began 1 second into a 5 second trial.
- •A probe sound occurred randomly between 1.3 and 3.8 seconds.
- •Subject tapped the touchscreen with their finger to indicate estimated position of the square at the time of the probe sound.
- •Eyelink 1000 eye-tracker, and ELO touchscreen recorded eye-movements and finger-tap response.

•4 main conditions:
- Occlusion present (A & B) vs. absent (C &D)
- Landmarks present (A & C) vs. absent (B & D).

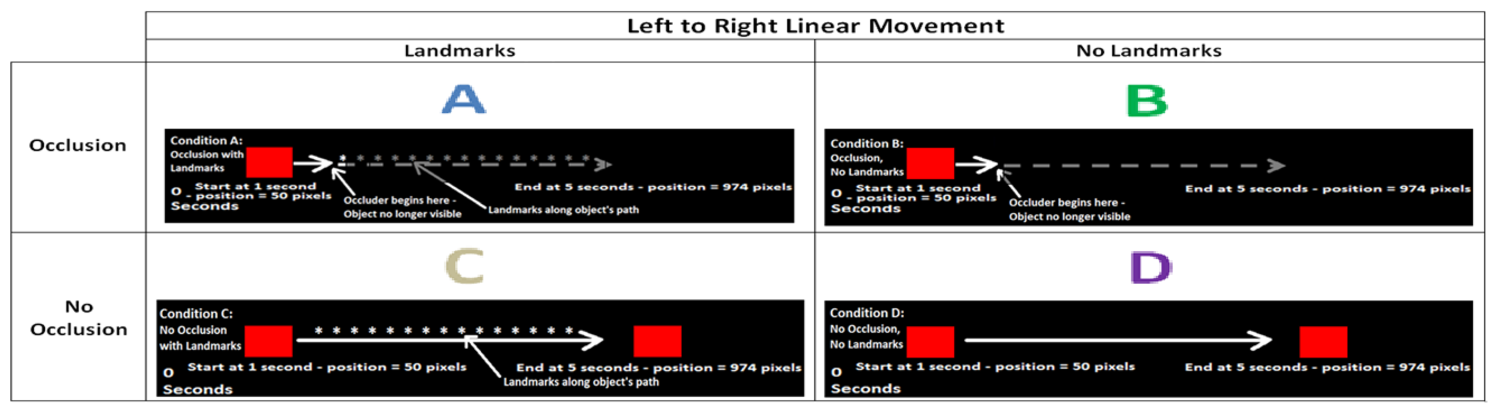

Figure 1: Illustration of object movement that subjects tracked with and without Occlusion and landmarks

Results

Sample trial of single object tracking (SOT) during disrupted viewing

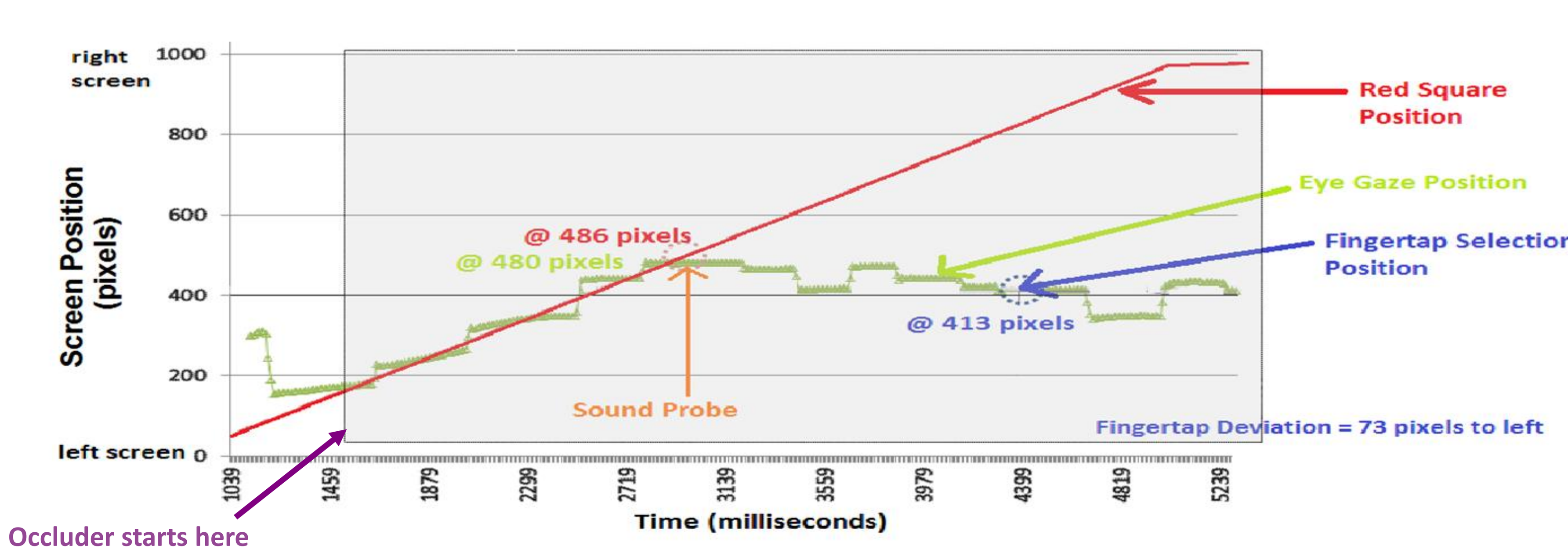

Figure 2: Sample of single object tracking trial from one subject showing horizontal gaze-position (green), finger tap selection (blue) and motion path of tracked object (red). This participants uses discrete eye-movements to accurately track the square up until sound probe (orange) when eyes stabilize on same position until tap response is made (about 1 second later). This sustained gaze-response may serve as a "place-holder" to help keep track of the position of the tracked object.

Results

Finger-Tap & Gaze Localization of (LR-moving) Tracked Object

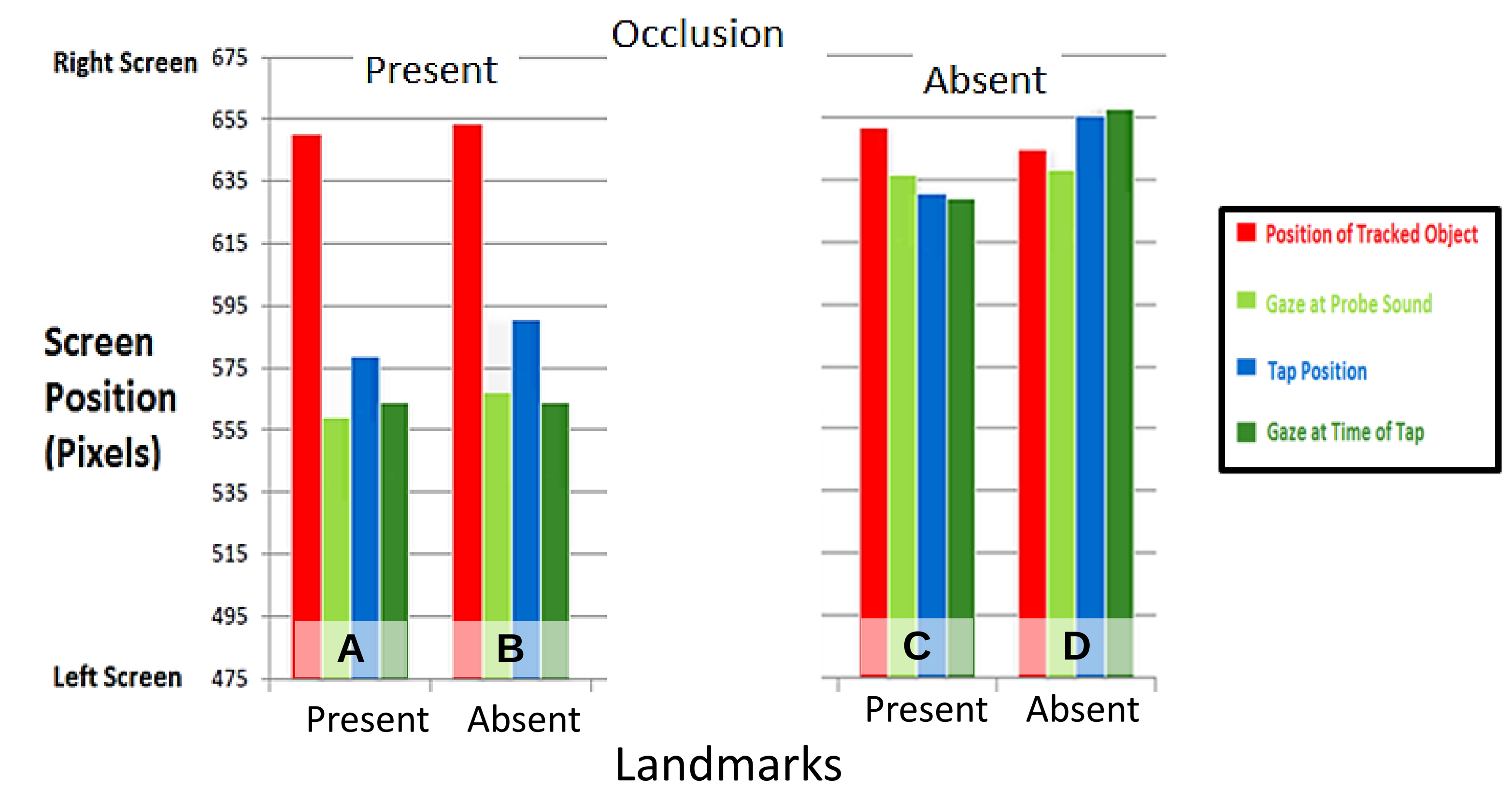

Figure 3. When estimating position of tracked-object at time of sound probe, gaze (green) and tap (blue) lag strongly behind hidden object (red) when occluded. Gaze-lag is slightly greater than tap-lag (p<.05). This is consistent with bias in eye-movement system which may drive lag-bias in tap response. Stable gaze-position at sound probe & later at tap selection (light vs. dark green) suggest eyes function as "place-holder" (as shown in Figure 1).

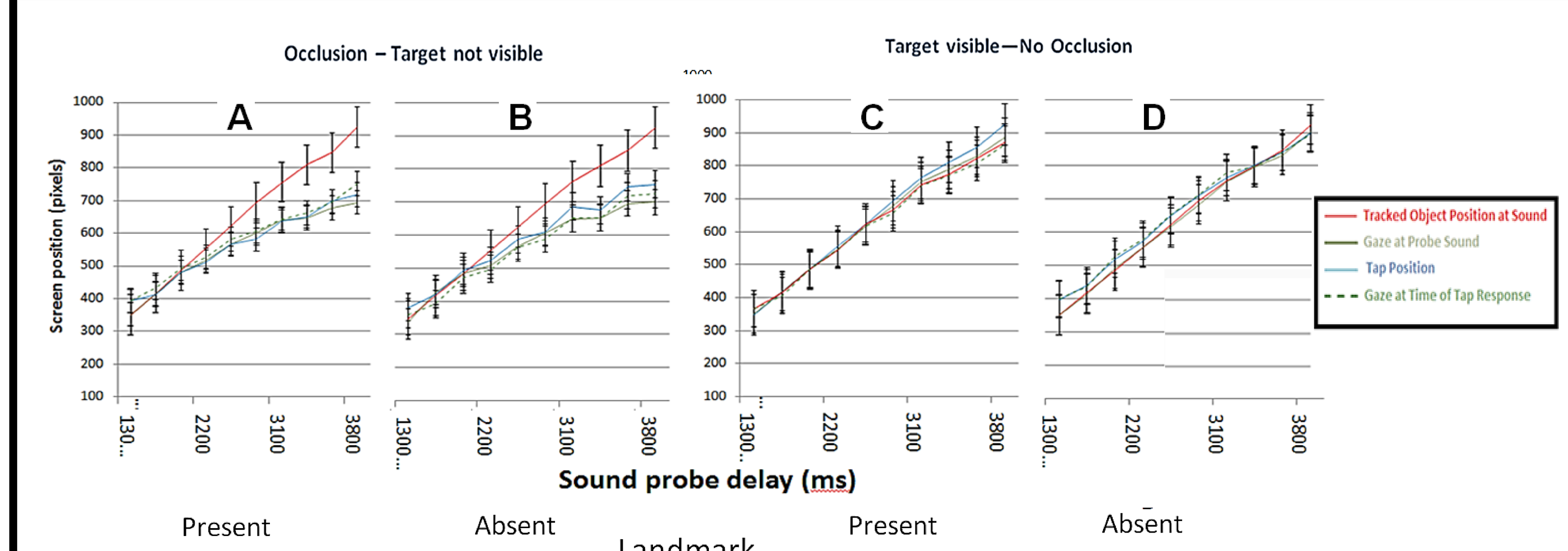

Figure 4A: When tracked object is occluded and landmarks are present (A), gaze and tap lag behind object position. As probe sound is delayed, lag increases, and gaze may lag slightly more than tap (consistent with our theory that gaze-bias may drive response-bias).

Figure 4B: Response lag disappears (C) or may reverse to slight overshoot (D-RL). Only a slight lag in the longest sound probe delays (LR-C), and when there are no landmarks (D). Most accurate localization in C suggests that landmarks may help overcome response bias but only when tracked object is visible.

Localization of Tracked Object

(Response - Object Deviation)

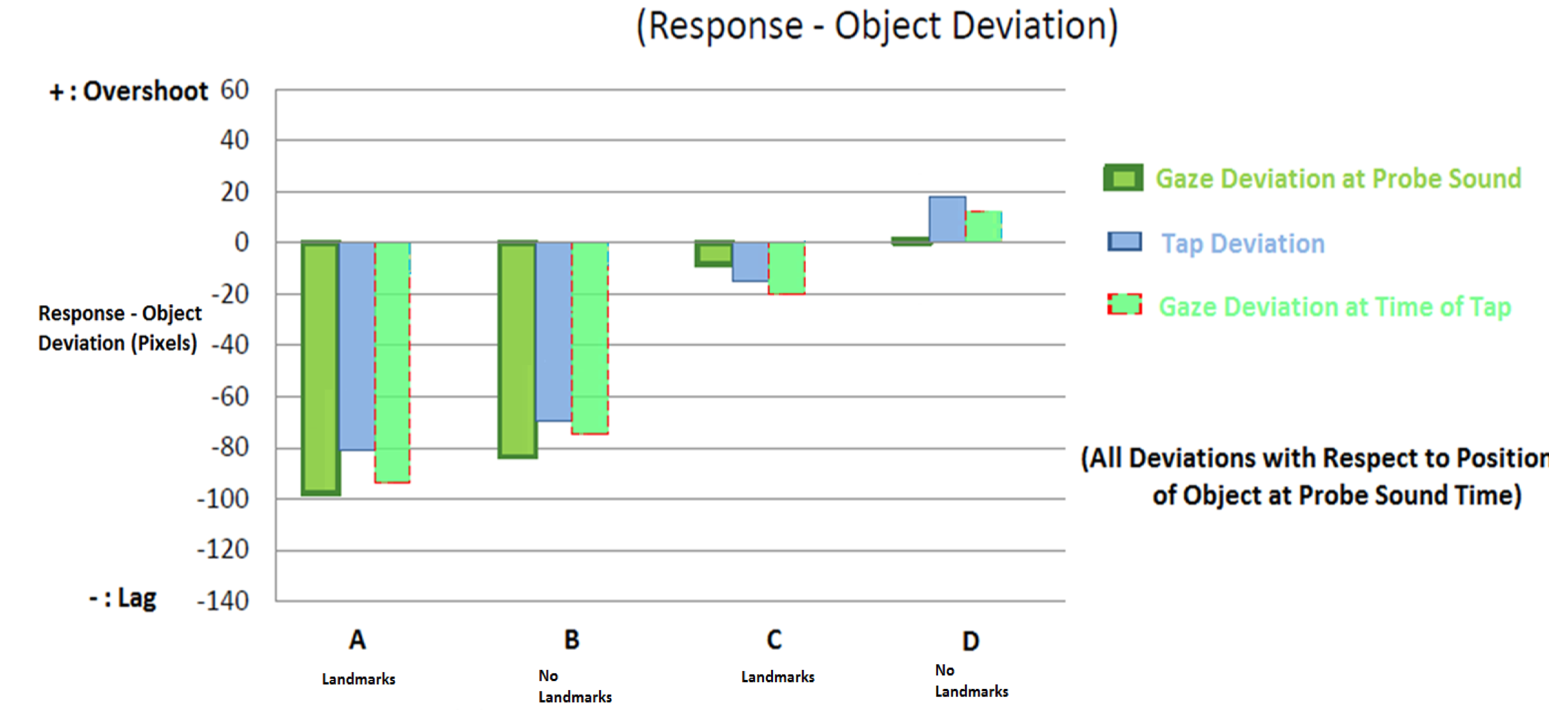

Figure 5: Deviations from tracked object with (object & response) direction standardized so that negative scores indicate lag, positive scores indicate overshoot and zero = perfect localization where response and objects positions are the same. Lag in tap (blue) and gaze (green) is much more pronounced in occlusion trials (A&B), with slightly greater lag for gaze at sound probe. These reflect same patterns described in Figures 1 to 4 (A&B). When the tracked object is visible, and landmarks are present, there is a slight lag for gaze & tap (C). When there are no landmarks, gaze at sound-probe is most accurate, but gaze & tap have slight overshoot at time of tap. Similar response deviations suggest behaviors are well-coordinated.

Summary

Effect of Occlusion on tracking:

Localization is impaired when tracking hidden objects. Similar to Pylyshyn & Cohen (1999), this argues against use of image-based mental mechanism resembling ordinary visual perception.

Effect of Landmarks on tracking:

There was a negligible effect of landmarks. Any benefit likely was overshadowed by ambient room lighting and incidental use of room reference frames.

"Lag" bias

- When tracking an occluded object, people underestimate its position along its motion path. This "lag bias" occurred for both gaze and finger-tap localization, and increases proportionally to the delay of sound probe.
- Gaze lags behind position of occluded object more than finger-tap. This may indicate a bias in eye-movement system driving lag in manual response & other estimates of object position.
- Similarity of gaze and finger- tap responses suggests these are well-coordinated behaviors.
- Because participants were instructed to indicate object localization with a finger tap, gaze is automatic. This suggests that a (spontaneous) gaze bias may drive the tap response, and perhaps also, the perceived location of the occluded object.

Eye as placeholder?

- Smooth pursuit eye-movements are used while tracking a visible object; when object is hidden eyes continue moving but in discrete saccades. At sound probe, eyes tend to halt and sustains the same approximate position until finger-tap selection is made.
- Our ongoing analysis is examining the reliability of sustained eye-movement from gaze-at-sound-probe to finger-tap selection, and to learn whether:
  - This pointer-like behavior of eye movements may reflect the use of FINST indexes to track the moving object (Pylyshyn, 2001).
  - There is a parallel to MOT findings of Keane & Pylyshyn (2006), and Aks et al (2010): Tracking improves when objects halt during their disappearance, and gaze often lags behind tracked objects. Perhaps, the placeholder & lag effect, and the halt advantage, are due to a common mechanism used in maintaining the position of occluded objects.

References

Aks, D., Haladjian, H., Kosmides, A., Annamraju, S., Kourtev, H. & Pylyshyn, Z. (2010). Blink-induced masking and its effect on Multiple Object Tracking: It's easier to track those that stop during interrupted viewing. *Journal of Vision*, 10(7),301a.

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.

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.

Pylyshyn, Z. W., & Cohen, J. (May 1999). Imagined extrapolation of uniform motion is not continuous. Paper presented at the Annual Conference of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, FL.

Pylyshyn, Z. W. (2007). Things and Places: How the mind connects with the world. Cambridge, MA: MIT Press.