Eye and Pointer Coordination in Search and Selection Tasks

Hans-Joachim Bieg*,
Lewis L. Chuang, Roland W. Fleming
Max Planck Institute for Biological
Cybernetics, Tübingen, Germany

Harald Reiterer
Human-Computer Interaction Group
University of Konstanz, Germany

Heinrich H. Bülthoff
Max Planck Institute for Biological
Cybernetics, Tübingen, Germany and
Dept. of Brain and Cognitive Engineering,
Korea University, Seoul, Korea

Abstract

Selecting a graphical item by pointing with a computer mouse is a ubiquitous task in many graphical user interfaces. Several techniques have been suggested to facilitate this task, for instance, by reducing the required movement distance. Here we measure the natural coordination of eye and mouse pointer control across several search and selection tasks. We find that users automatically minimize the distance to likely targets in an intelligent, task dependent way. When target location is highly predictable, top-down knowledge can enable users to initiate pointer movements prior to target fixation. These findings question the utility of existing assistive pointing techniques and suggest that alternative approaches might be more effective.

CR Categories: H.5.2 [Information Interfaces and Presentation] User Interfaces – Input devices and strategies. H.1.2 [Models and Principles] User/Machine Systems – Human factors

Keywords: Eye movements, Eye-tracking, Eye-hand Coordination, Multimodal Interfaces, Input Devices

1 Introduction

A fundamental task in modern computer interfaces is the selection of a graphical item with a pointing device (e.g. a mouse). The difficulty of such a pointing task is well described by Fitts' Law as the logarithm of the amplitude (distance to the target) of the movement divided by the target width (the required accuracy) [MacKenzie 1992]. Several techniques have been developed to facilitate pointing tasks by decreasing their difficulty along the lines of Fitts' Law [Balakrishnan 2003]. For instance, techniques exploit eye-tracking technology to move the pointer to the location of a detected eye fixation on the computer screen [Zhai, et al. 1999; Bieg et al. 2009; Drewes and Schmidt 2009] or to select one of several pointers near the fixation point [Räihä and Špakov 2009; Blanch and Ortega 2009], thereby reducing the amplitude and thus also the difficulty of the pointing movement. The underlying assumption is that the item of interest is fixated before it is acquired and that turn-taking occurs between eye and pointer movement to maximize movement time savings. The purpose of this work is to test directly whether this assumption holds across a range of tasks.

*e-mail: hans-joachim.bieg@tuebingen.mpg.de

Eye and hand coordination has been extensively studied for reaching and pointing actions with the hands [e.g. Neggers and Bekkering 2000] but not for tasks with a computer input device in a graphical user interface. A first study was carried out by Smith et al. [2000] to examine differences in the coordination of eyes and pointer movements when using various computer input devices. Smith and colleagues found that participants employed three different eye and hand coordination strategies when carrying out simple target acquisition tasks: a) eyes leading the pointer to the target (target gaze), b) eyes directly tracking the pointer (following), and c) switching between target and pointer (switching). Similar patterns were found by Law et al. [2004] albeit with a specialized laparoscopic tool and not with a common computer pointing device.

In the present study we investigated the coordination of eyes and pointer movements on several selection tasks. In contrast to previous studies, these tasks required visual search for the target item before selection. Such a combination of search and selection is a staple of numerous human-computer interaction tasks [Fleetwood and Byrne 2008]. We wanted to see whether and how visual search affected the coordination of eye and mouse pointer movement.

2 Method

2.1 Participants

The experiment was performed on 12 participants with normal or corrected-to-normal vision. Data from one participant was removed because of large eye-tracker calibration errors. All participants were regular computer and mouse users.

2.2 Tasks

Three tasks were presented (T1-T3) that required searching for and selecting (i.e. clicking on) one or several target items or drag-and-drop of the target item (see Figure 1). A fourth task was also presented during the experiment but is not discussed here. Task T1 was a simple target acquisition task. Users started at a location 800 px to the right of the target item and moved a graphical pointer with the mouse to select it and drag it over to the circular area. Task completion was confirmed by clicking on a button labeled "done" next to the circular area. In the second task (T2) participants searched a 13 × 13 grid for a single target, specified uniquely by shape (triangle, rectangle, circle) and color (red, green, blue) before performing the same drag-and-drop task. The position of the grid on the screen was fixed and the locations of target and distractor items were randomized within the grid. In task T3, the same amount of items were scattered across the whole screen. Five target items had to be searched and clicked on in this task. The size of each item was approximately 20 px (3.3 cm, 1.92° visual angle). Before each trial, a

screen was shown to inform participants of the current target's color and shape. For reasons of consistency, this was also done in task T1. This preparatory screen contained an instance of the target item and an "ok" button which had to be clicked in order to start the trial. For tasks T1 and T2, the button and item instance were located at a constant position to the right of the screen. The button was at the exact same position as the "done" button in the trial display. The item instance was above the button. For task T3, item and button were shown in the center of the screen. With this arrangement, gaze and pointer position were at an approximate constant starting position for each trial (near the "done" button on the right side for T1 and T2 or center of the screen for T3).
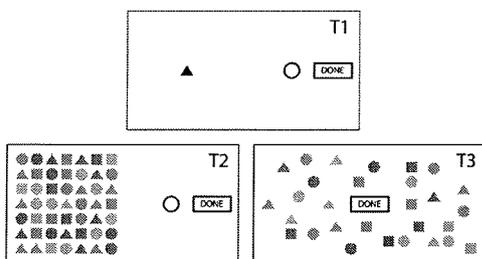

Figure 1 *Schematic representation of the experimental tasks.*

2.3 Apparatus

Participants used an optical computer mouse to complete the task (Logitech Optical Wireless). The experimental computer ran Ubuntu Linux version 8.10. Participants sat 100 cm in front of a large back-projected screen which displayed the graphical stimuli at a rate of 120 Hz (resolution 1400 × 1050 px, size 94.1° × 89.1° visual angle, 2,15 × 1,97 m). While carrying out the tasks, participants' gaze was tracked with a mobile eye-tracking system. The system allowed for unconstrained head motion by combining a head mounted eye-tracker (SR Research Eyelink II) with an infrared motion capturing system (Vicon MX 13 with four cameras) using the libGaze tracking library [Herholz et al. 2008]. With this system, gaze was computed in terms of screen coordinates at 120 Hz.

2.4 Procedure

Each experimental session lasted around 90 minutes with short pauses every 15 minutes and regular recalibration of the eye-tracker before each block. Prior to data collection, the tasks were explained, demonstrated to the participants, and one round of practice trials was presented. Each participant then carried out five blocks of tasks. Each block started with 10 trials of task T1, followed by 10 trials of T2, and 3 trials of task T3.

3 Results

It is well known from experiments on visual search that search for combinations of features (e.g. shape and color) take more time as a function of the number of items and difficulty of discrimination. In addition, the number of fixations is correlated with search difficulty [Zelinsky and Sheinberg 1997]. We compared trial completion times and number of fixations of tasks T1 and T2 to see whether an effect of the added requirement to search an item prior to selection was present. The mean trial completion time for a simple drag-and-drop trial (T1) was 3.5 s (std. dev.: 0.45). In the search condition (T2) completing the

trial took considerably longer (mean: 5.5 s, std. dev.: 0.60, 95% CI of difference: 1.5-2.3 s, t(10)=10.5, p<0.01). Fewer fixations were made during the single-target task T1 (mean: 6.0, std. dev.: 0.78) compared to the search task T2 (mean: 14.4, std. dev.: 2.37, 95% CI of difference: 6.7-10.0, t(10)=11.1, p<0.01). Thus, it took participants approximately 2 seconds longer to complete task T2. This time was likely spent on the search for the target item and eye movements required for item identification.

Since almost 40% of trial completion time seems to be spent on search, how do users coordinate pointing movements during this time? Analysis of pointer movement in task T2 shows that the pointer was already close to the target when the first fixation on the target item was made. For instance, in more than 80% of trials, the pointer had already moved half of its initial distance to the target (Figure 2 line a). For the simple selection task T1, the pointer's distance to the target was generally larger. Still, in almost 60% of trials, the pointer had already moved more than half of the initial distance (Figure 2 line b). In task T2, the median latency between the first fixation on the target item and the beginning of the pointer movement was -1.7 s. In T1, the median latency was -190 ms. In more than 60% of trials of T2 the pointer movement was initiated almost 1 s before the target had been fixated (Figure 3 line a). In trials of T1, the latency was shorter but still mostly negative.
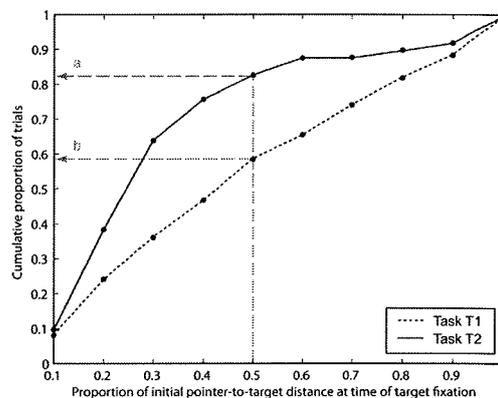

Figure 2 *Cumulative histogram of the distance (as proportions of the initial pointer-to-target distance of approximately 800 px, 68° visual angle) between the pointer and the target item at the time of the first fixation on the target.*

Two explanations for these findings seem probable. First, participants moved the pointer immediately once the trial started by using knowledge from previous trials about the approximate target location (T1) or grid location (T2). This movement required no visual guidance and was carried out while the eye made a saccade to the target to guide the final targeting phase of the movement. Similar findings were also reported for hand movements by van Donkelaar et al. [2000].

A second explanation could be that, in the case of T2, participants parallelized search and pointer movement. To investigate this, we analyzed pointer movements during task T3. In this task target locations could not be predicted from memory. A typical example of eye scanpaths and pointer motion is shown in Figure 5. Trajectories for the vertical and horizontal dimension along the time axis are shown in Figure 5 b and c. From the inspection of individual trial data it seems that the pointer roughly followed the scanpath throughout the trial. To test this hypothesis, we

calculated the mean Euclidean distance between the location of the mouse pointer and eye gaze across all trials of T3 for all participants. The mean Euclidean distance was minimal when the mouse pointer signal was shifted a few hundred milliseconds forward in time (see Figure 4). Even though the extent of the shift varied between participants (see participant means in Figure 4) it was positive for all participants. A positive shift can be conceived as a shift of the mouse pointer path to the left in Figure 5 b and c. This means the mouse pointer generally lagged behind the eye movements but it does not give any clues to whether it did so consistently throughout the trial.
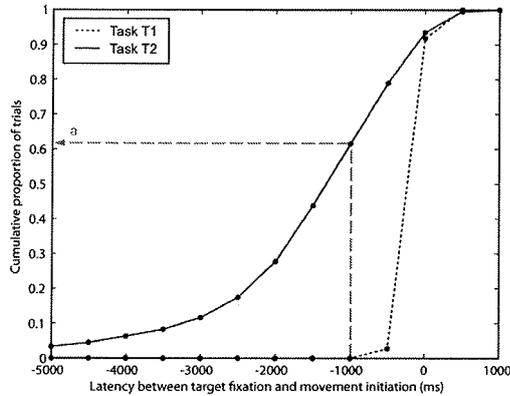

Figure 3 *Cumulative histogram of the latency between a fixation on the target and movement initiation (the point in time when the pointer was moved more than 40 px from the starting position).*
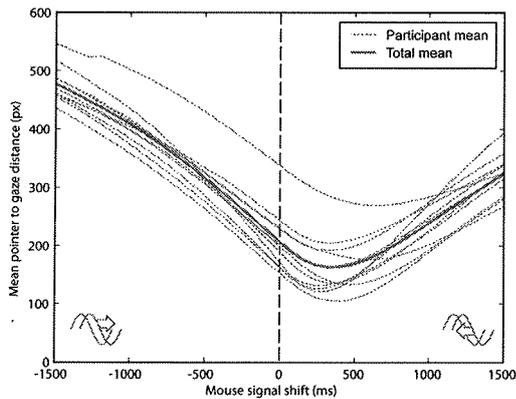

Figure 4 *Mean distance of the pointer to the gaze point on the screen across trials of task T3 as a function of the amount of phase shift of the mouse pointer "signal". Pictograms indicate effect of shift.*

To analyze whether the pointer followed gaze throughout a trial, we implemented a model of pointer motion in which the pointer was moved directly from selection point to selection point and remained motionless during the periods in between. This model simulates the turn-taking strategy assumed by some assistive interaction techniques in which a user moves the mouse to acquire a target but keeps it stationary while searching. Compared against actual pointer motion the mean Euclidean distance across all trials and participants was significantly larger for this turn-taking model (95% CI difference: 238-289 px, $t(10)=23.0$,

$p<0.01$). This indicates that participants generally moved the pointer while searching and did not adopt a turn-taking strategy.
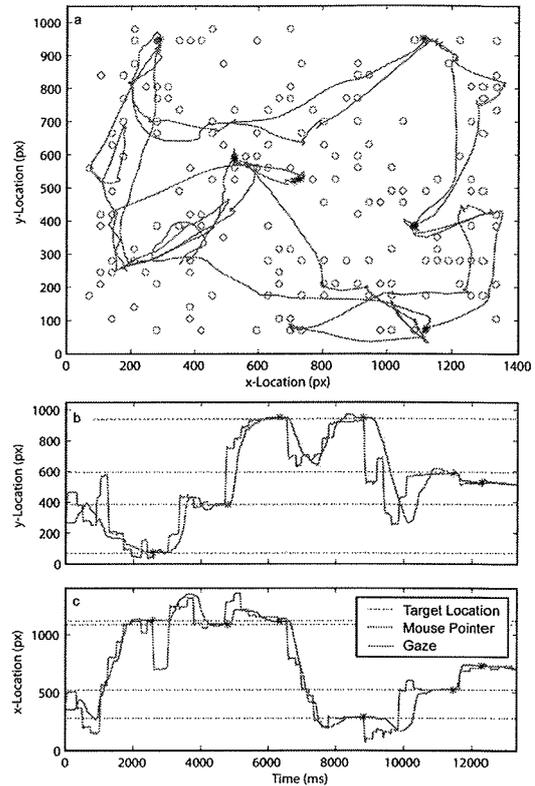

Figure 5 *Typical eye scanpath and path of pointer motion for a single trial of task T3. Item locations are shown as gray circles (color/shape coding not shown) and target items are circles colored in red. Selection actions are depicted as blue stars.*

4   Conclusion

In this study, we investigated the coordination of eye movements and mouse pointer movements during a range of search and selection tasks. In performing these tasks, participants use two main eye-hand coordination strategies. First, when the approximate location of the target is known, participants perform pointer movements that are initiated without eye guidance, possibly in parallel to an eye movement to the target item. Second, in tasks requiring visual search for a target item prior to selection, users parallelize search and pointer movements. This behavior might have two functions: First, in order to utilize "slack motor time" while looking for the target, users roughly follow their scanpath to keep the pointer close to the searched area. Thereby they minimize the amplitude of the acquisition movement, once a target is found. Second, users might employ the pointer as an active reference tool for marking potential target items or as a reference for keeping track of the search area. Such "deictic" pointing behavior [Ballard et al. 1997] has also been observed in menu selection and web browsing tasks [Cox and Silva, 2006; Rodden et al. 2008]. Further analyses and experiments are required to clarify the function of the observed behavior in the current task.

Our results also show that the eye reaches the targeted item prior to the pointer. This is in line with the "target gaze" behavior identified by previous studies [Smith et al. 2000; Law et al. 2004]. However, our findings also imply that the eyes fixate on the target rather late in the pointing process if the approximate target location is known. This gives rise to the assumption that eye fixations are only needed for precise targeting motions during the final pointing phase. We did not find evidence for the "following" or "switching" behaviors described by Smith and colleagues [2000]. We cannot exclude the possibility that these behaviors may occur. In the context of search tasks, however, our data suggest that participants consistently follow their gaze with the mouse pointer rather than the other way round.

These findings also have implications for the design of optimized pointing techniques, for instance, gaze-assisted techniques that reduce the amplitude of the pointing movement [e.g. Zhai et al. 1999]. An implicit assumption of these techniques is that the eyes precede a pointer movement so that movement time savings could be gained by moving the pointer to the point of gaze. Results of the present study indicate that users tend to minimize pointing amplitudes with little effort by carrying out approximate pointer motions in parallel to visual search when the target is unknown or initiating movements even before fixating the target when the target location is known. Both scenarios are very common in many user interface tasks. Locations of interface elements such as window controls are often well known whereas content elements such as links on a web page must often be searched [Rodden et al. 2008]. In fact it is hard to construct a scenario where no knowledge is available or no search is needed. It might therefore be reasonable to explore other optimization techniques, for instance, techniques that do not minimize movement amplitudes but the needed endpoint precision, for instance using target-expansion [e.g. McGuffin and Balakrishnan 2002].

5 Acknowledgements

References

BALAKRISHNAN, R. 2004. "Beating" Fitts' law: virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, 61, 857-874.

BALLARD, D. H., HAYHOE, M. M., POOK, P. K., AND RAO, R. P. N. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20, 723-767.

BIEG, H., CHUANG, L. L., AND REITERER, H. 2009. Gaze-Assisted Pointing for Wall-Sized Displays. In *Human-Computer Interaction – INTERACT 2009 (Proceedings of 12th IFIP TC 13 International Conference)*. Springer, 9-12.

BLANCH, R., AND ORTEGA, M. 2009. Rake cursor: improving pointing performance with concurrent input channels. In *Proceedings of CHI*. ACM, 1415-1418.

COX, A. L., AND SILVA, M. M. 2006. The role of mouse movements in interactive search. In *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*. Lawrence Erlbaum Associates, 1156-1161.

DREWES, H., AND SCHMIDT, A. 2009. The MAGIC Touch: Combining MAGIC-Pointing with a Touch-Sensitive Mouse. In *Human-Computer Interaction – INTERACT 2009 (Proceedings of 12th IFIP TC 13 International Conference)*. Springer, 415-428.

FLEETWOOD, M. D., AND BYRNE, M. D. 2008. Modeling the visual search of displays: a revised ACT-R model of icon search based on eye-tracking data. *Human-Computer Interaction*, 21, 2, 153-197.

HERHOLZ, S., CHUANG, L. L., TANNER, T. G., BÜLTHOFF, H. H., AND FLEMING, R.W. 2008. LibGaze Real-time gaze-tracking of freely moving observers for wall-sized displays. In *Proceedings of the Vision, Modeling, and Visualization Workshop (VMV)*. IOS Press, 101-110.

LAW, B., ATKINS, S. M., KIRKPATRICK, A. E., LOMAX, A. J., AND MACKENZIE, C. L. 2004. Eye Gaze Patterns Differentiate Novice and Experts in a Virtual Laparoscopic Surgery Training Environment. In *Proceedings of the 2004 symposium on eye tracking research & applications*. ACM, 41-47.

MACKENZIE, S. I. 1992. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction*, 7, 1, 91-139.

McGUFFIN, M., AND BALAKRISHNAN, R. 2002. Acquisition of expanding targets. In *Proceedings of CHI*, ACM Press, 57-64.

NEGGERS, S. F. W., AND BEKKERING, H. 2000. Ocular Gaze is Anchored to the Target of an Ongoing Pointing Movement. *Journal of Neurophysiology*, 83, 639-651.

RÄIHÄ, K.-J., AND ŠPAKOV, O. 2009. Disambiguating Ninja Cursors with Eye Gaze. In *Proceedings of CHI*. ACM, 1411-1414.

RODDEN, K., FU, X., AULA, A., AND SPIRO, I. 2008. Eye-Mouse Coordination Patterns on Web Search Results Pages. In *Proceedings of CHI extended abstracts*. ACM, 2997-3002.

SMITH, B. A., HO, J., ARK, W., AND ZHAI, S. 2000. Hand Eye Coordination Patterns in Target Selection. In *Proceedings of the 2000 symposium on eye tracking research & applications*. ACM, 117-122.

VAN DONKELAAR, P., AND STAUB, J. 2000. Eye-hand coordination to visual versus remembered targets. *Experimental Brain Research*, 133, 414-418.

ZELINSKY, G. J., AND SHEINBERG, D. L. 1997. Eye Movements During Parallel-Serial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1, 244-262.

ZHAI, S., MORIMOTO, C., IHDE, S. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. *Proceedings of CHI*. ACM, 246 – 253.