

Gaze-based Cursor Control Impairs Performance in Divided Attention*

Róbert Adrian Rill^{ab} and Kinga Bettina Faragó^a

Abstract

In this work we investigate the effects of switching from mouse cursor control to gaze-based control in a computerized divided attention game. We conducted experiments with nine participants performing a task that requires continuous focused concentration and frequent shifts of attention. Despite carefully controlling experimental and design aspects, the performance of subjects was considerably impaired when using gaze-based control. The participants were experienced users of the mouse control version of the task, we adjusted the difficulty to the more demanding conditions and selected the parameters of gaze input based on previous research findings. In contrast to our assumptions, experienced users could not get used to gaze-based control in the amount of experiments we performed. Additionally we consider the strategies of users, i.e. their method of problem solving, and found that it is possible to make progress in our task even during a short amount of practice. The results of this study provide evidence that the adoption of interfaces controlled by human eye-gaze in cognitively demanding environments require careful design, proper testing and sufficient user training.

Keywords: gaze-based control, eye tracking, divided attention, human performance, cognitive load, Midas Touch, dwell time

1 Introduction

The seminal work of Yarbus [30] had an enormous impact on consequent researches on eye movements. He showed that it is possible to infer the task an observer is performing from their fixation patterns. Not only do eye movements indicate the search target during visual exploration, but they also reveal emotions, intentions

*This research has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013).

^aELTE Eötvös Loránd University, Budapest, Hungary. Faculty of Informatics, 3in Research Group, Martonvásár, Hungary.
E-mail: rillroberto88@yahoo.com (ORCID id: <https://orcid.org/0000-0002-3004-7294>), kinga.farago@gmail.com (ORCID id: <https://orcid.org/0000-0002-2402-7123>)

^bFaculty of Mathematics and Computer Science, Babeş-Bolyai University, Cluj-Napoca, Romania.

and cognitive processes [2]. Eye contact and gaze direction represent a powerful means of communication in regulating interaction and establishing socio-emotional connection [17].

In addition to having a long history in studying human visual behavior in medical and psychological research, eye tracking has the potential to revolutionize interface design in the field of human-computer interaction. Due to the technological advancements, non-intrusive and accurate hardware solutions are readily available (see, e.g., [8, 17]). Moreover, explosion in computer vision in recent years lead to the development of state-of-the-art appearance-based gaze estimation methods that use only images and videos from off-the-shelf monocular RGB cameras, and have an acceptable range of errors in prediction (see, e.g. [22] and the references therein).

Eye movements can be divided into several different categories [5, 12, 16, 17]. Saccades are the most common way of moving the eyes in a sudden, ballistic way of 2 degrees or larger taking about 30-120 ms each. They are typically followed by fixations of at least 100 ms, generally between 200-600 ms periods of relative stability. However, during fixations the eyes still make small, jittery motions, covering usually less than one degree. Blinks of up to 200 ms may occur during a fixation without terminating it. Smooth pursuit movements are less sudden than saccades and occur only in response to a moving target in the field of view. There are also other eye movements which, however, are not significant in human-computer interaction. For more information about eye movements, see [17] and the references therein.

Modern technological solutions facilitate the design of complex human-computer interaction interfaces. These often require the divided attention of users. Compared to the slow and deliberate way of operating a mouse or other input device, eye movements usually scan the screen involuntarily, for example the user is not aware of the jittery motions during a fixation. Moreover, eyes are used primarily for perception [18] and they typically precede actions [1, 14, 17, 27]. Thus, the additional use for control requires careful design of human-computer interaction interfaces [3, 13, 20] in order to provide adequate feedback and to avoid false activation.

Controlling computers by gaze could also improve performance in multitasking situations by taking advantage of the nature of eye movements, reducing this way cognitive load or increasing safety. For instance, gaze-based control of secondary displays in automotive or aviation environments can reduce pointing and selection times while also performing the primary task of driving and flying, respectively [23].

1.1 Related work

Despite the problems of using the same modality for both perception and control, gaze estimation and eye tracking have important application areas ranging from medical diagnosis and psychological research to the design of interfaces and usability studies of gaze-controlled applications from the field of human-computer interaction [17]. Probably the most common example is eye-typing [15, 16, 18,

19, 25, 28]. Other applications include object selection on interfaces [12] and in real world [26], target search and selection [28], computer game control (ranging from simple puzzles [1] and classical games [6] to role-playing and first-person shooter video games [11, 27]), facilitating attention switching between multiple visual displays [13], robotic device control [7], web-browsing [5, 14], interacting with geographic information systems [3], developing interactive graphical user interface elements [20], projected display control in automotive and military aviation environments [23].

Using gaze as an input method may not be comparable to the universal mouse and keyboard [14] because of the nature of the human eye movements and the physiology of the eye [17], yet it can still have several advantages. For people with physical disabilities gaze-based interaction provides a means to communicate and interact with technology and other people [5, 6, 17, 18, 20]. The performance of children with disabilities can be enhanced considerably through gaze controlled computers [10]. In the case of older adults, it may be able to compensate for the declined motor functions when using mouse input [21].

Another relevant factor to consider is the engagement of users. Eye movements are extremely fast and require little effort [1, 3, 7, 14, 25, 27]. Thus, gaze response may be well suited for visual search tasks and novice users may find captivating the natural quality of selection by looking [28]. Gaze can represent a superior input modality in simple computer games in terms of achievements, engagement or gameplay experience [1, 27], even for users without any previous training [6]. Furthermore, gaze-based interaction facilitates attention switching when using multiple displays [13], and eye movements represent a means of communication in collaborative virtual environments too [6, 27].

In contrast, controlling by gaze also faces considerable challenges [17]. Probably the most common issue is the *Midas Touch* problem [1, 3, 4, 5, 7, 12, 14, 17, 18, 20, 23, 26, 27], i.e. activating a command involuntarily only by looking at a specific target. Eye movements are largely automatic and unconscious. Normal visual perception requires that the eyes of the user scan the scene, gather information about the environment before making an action [1, 14, 17, 27]. Ideally the system should distinguish casual viewing from intentional control.

When using gaze as input, there is no natural counterpart of a mouse click [6, 12, 20]. For systems using solely gaze-based control for selection the most obvious and common alternative is *dwelling time* [1, 4, 5, 11, 12, 14, 15, 16, 17, 18, 19, 28], i.e. gazing at a specific object for a sufficiently long time to trigger an activation command. However, this also raises further questions. Dwelling time must be short enough to be comfortable for the subject, but this brings up the *Midas Touch* problem. On the other hand, a long dwelling time might ensure that unintentional selections are not made, but it limits exploration time, diminishes the advantage of fast and natural eye movements and reduces the responsiveness of an interface.

Other alternatives for dwell-based selections to avoid unintentional commands are gaze-gestures [3, 11, 17, 20, 26], predefined gaze patterns as pre-programmed strategies for control [7] or combining gaze-based search and pointing with other modalities such as additional hardware button (cf. [14]) or feet [3].

Investigating feedback modalities is also in the focus of researchers. When using dwell-time, the user only initiates the action and it is the responsibility of the system to provide a clear indication of the status [17]. Majaranta et al. [19] compared auditory and visual feedback during eye-typing. Their results show that the feedback method influences both text entry speed and error rate, and suggest that a simple auditory confirmation of selection is more effective than visual feedback alone. In their follow-up works the authors also compare long versus short [16] as well as adjustable [15] dwell time duration. In a more recent work Majaranta et al. [18] also found haptic feedback to produce results close to those of auditory feedback. Other feedback methods to consider are using animation to indicate the progression of dwell time [15, 18], zooming into the area of focus [13].

The state-of-the-art, low-cost and easily accessible eye tracking technology makes gaze input a useful, fast and convenient way of communication. Although a considerable amount of work has already been done on interaction techniques, there is still no general procedure on how to incorporate eye-movements into human-computer interaction interfaces in a natural and unobtrusive way. Furthermore, little research has been conducted to examine the effects of different input modalities on users' problem solving strategies. Present-day computer interfaces require users to employ a range of complex strategies, including planning, goal searching, handling interruptions and information coordination. Bednarik et al. [1] compared dwell time, gaze-augmented interaction and mouse input using a simple puzzle game and found that the interaction methods affected performance, problem solving strategies and user experience. Dorr et al. [6] showed that gaze is superior to mouse input in a classic computer game and found that expert and novice players differ in their employed eye movement strategies.

Investigating people's problem solving strategies, using gaze as input and comparing it to more classical response methods also represents a powerful tool in psychological research [7, 21, 25, 27, 28] (e.g., in visual search tasks), because it may reveal new aspects of cognitive processes and may have implications on the design process of interfaces employing gaze tracking. Therefore we investigate gaze-based control in a special dynamic divided attention task. Particularly, we designed and implemented a simplified version of the popular Train of Thought game from the Lumosity¹ online platform. Lumosity is comprised of a set of computerized games designed by scientists, each aiming to train one of five core cognitive abilities: attention, processing speed, memory, flexibility and problem solving [9]. In the following we will refer to our version of Lumosity's Train of Thought game as the *Divided Attention (DA) Game*.

The contributions of this paper consist in investigating the effects of switching from the traditional mouse to gaze-based input in a divided attention task not examined by previous works to the best of our knowledge. Despite carefully considering several experimental and design aspects, the performance of participants was considerably impaired by gaze-based control in this cognitively demanding task requiring the divided attention of players. The subjects in our experiments were

¹<http://www.lumosity.com/>

experienced in the conventional mouse control version of the game, we adjusted the difficulty of the task to the more demanding conditions and chose dwell-time and other parameters based on previous research findings.

The paper is organized as follows. In Section 2 we describe briefly the DA Game, detail our design choices, present the experiments with gaze-based control and define the performance measures. Section 3 presents the results of the experimental and statistical analysis. This is followed by a discussion in Section 4, which highlights future directions as well. Finally, Section 5 concludes the paper.

2 Methods

2.1 Design of task

In our previous work [24] we conducted a longitudinal study with mouse control and presented in detail the design process of the DA Game used in our experiments. Thus, here we only describe briefly the purpose of the task and present the design elements of the gaze-based control version.

The DA Game tests the divided attention and working memory of the players by requiring them to continuously focus on multiple simultaneous targets, to switch frequently between them keeping track of each one. The task of the user is to direct continuously oncoming objects to their color-matching destination through selecting and flipping switches at forks and changing this way the direction of the tracks and the path of the moving objects. For a snapshot of the game see Figure 1.

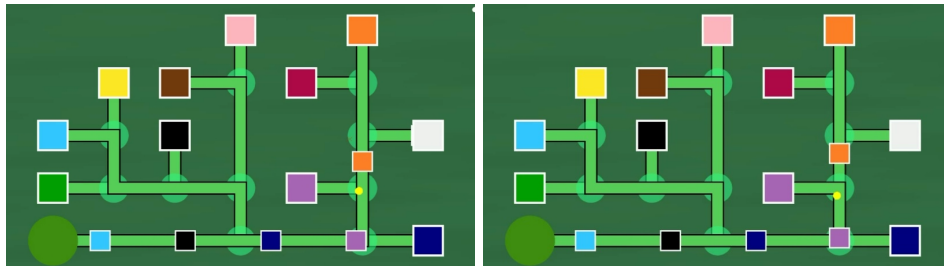


Figure 1: Snapshots of two frames from the Divided Attention Game used in our experiments: the small squares are moving continuously and have to be directed to their color-matching destinations by flipping switch nodes represented by transparent green circles. The yellow dot corresponds to the screen coordinates of the user's gaze direction. The player started to fixate on the switch next to the purple destination on the left image; the predefined dwell time of 500 ms has just elapsed and the switch was flipped as seen on the image on the right.

The traditional mouse cursor was replaced by a yellow dot displayed at the screen coordinates of the gaze direction. Although in some cases it might distract users' attention, due to the nature of eye movements, we decided to show the cursor

at all times in order to provide continuous feedback for the players in a task with time constraints and to allow the possibility to compensate for potential drifts of the eye tracker during one gameplay. The noise of the device and the jittery movements of the eyes during fixations may also distract the concentration of users [17] when performing the task. It is easier to keep a steady cursor in one place until the target is selected. Accordingly, we applied a smoothing to the cursor movement using a moving average window on each of 5 consecutive samples. This does not slow down the responsiveness in the DA Game, which requires fast-paced user actions.

The most essential part of gaze-based control interfaces is generating a selection, i.e. flipping a switch node in the DA Game in our case. We chose the most common method, namely dwell time. We selected the length of the interval based on previous works enumerated next. We note that after the dwell time elapsed, we did not give any additional feedback to the user about the fact that the switch was flipped, since this is clearly noticeable as seen on Figure 1. Also, in order to be consistent with mouse-based selection, if the user continued to fixate on the target, it was selected repeatedly when the dwell time had elapsed again.

2.1.1 Dwell time duration

Jacob [12] found that a short dwell time of 150-250 ms gave excellent results, while duration over 750 ms was not useful at all in object selection tasks. In [28] the authors state that fixations longer than 500 ms are often seen during cognitive integration phases of difficult tasks. Their pilot studies indicated that 700 ms or less works well for simple tasks. They also found that a dwell time of 1000 ms makes false selections unlikely in a target selection task and that 750 ms is subjectively slow in their eye-typing task.

Majaranta et al. compared short and long dwell time duration, i.e. 450 ms vs 900 ms [16, 19]. Experienced participants achieved faster typing speed but higher overall error rates. The authors concluded that with short dwell time sharp and clear feedback is essential. In a later work, Majaranta et al. [15] also investigated adjustable dwell time in a longitudinal study and found that dwell time decreased from an average of 876 ms to 282 ms, and error rates also decreased. It is important to note that the learning rate was rapid during the first few sessions and decelerated prominently. More recently Majaranta et al. [18] found a dwell time of 860 ms in a practice session too long, and 500 ms seemed to work for them. They also mention that 500 ms might be too fast for novices in eye-typing. Expert typists may even use dwell times that correspond to their normal fixation times (for more details see [18] and the references therein).

Kern et al. [13] used a delay of 600 ms for marking gaze positions to reduce attention switching costs between multiple computer screens. Hyrskykari et al. [11] used a dwell time of 700 ms in a multi-user role-playing game where the user's gaze has to be maintained in the center of the screen for most of the time. Fedorova et al. [7] employed fixations of 500 and 300 ms long for robot control. They note that this resulted in slow but reliable communication, in situations where distractors are common and false alarms can have high costs. Lutteroth et al. [14]

used only a 200 ms activation dwell threshold in a web browsing task and achieved a fairly close performance to the mouse click alternative. Chen and Shi [5] investigated variable dwell time in a web-browsing task using probabilistic models and their best model reduced error rate by 50% and response time by 60% while maintaining the other performance measure constant when compared to a uniform dwell time of 100 and 300 ms, respectively. They also used in their practice experiments a fixed 500 ms dwell time.

Based on the above studies and also taking into consideration the fact that visual reaction time is considerably less than 500 ms [29], we selected a dwell time duration of 500 ms.

2.2 Participants and experiments

In our previous work [24], we have performed a longitudinal study with 10 participants, who were asked to play with the regular mouse control version of the DA Game. The volunteers were aged between 25 and 30 (mean age was 27 years, SD=1.76), had normal or corrected-to-normal vision and reported no attention disorders nor color vision deficiency. The experiments lasted several days, with multiple trials played each day. We manipulated the difficulty of the game, i.e. the moving speed of the squares, according to the score of the players from the previous trial. Based on this, the experiments were separated into three phases: beginner, intermediate and advanced.

For the experiments with the gaze-based control version, 9 out of the 10 participants were invited back for ten additional trials. The participants were instructed that data about their gameplays will be logged for further analysis and they were asked to sign a consent form before the experiments. We also allowed rest periods after each trial if the subject requested so.

For gaze tracking the Tobii EyeX Controller² [8] device was used, which is attached to the bottom of the display, has a sampling rate of 60 Hz and requires personal calibration before each data collecting session. Although the manufacturers claim that no continuous recalibration is required [8], drifts may occur over time [28] due to illumination or head position changes. Accordingly, we repeated the calibration procedure between trials when necessary.

2.3 Performance measures

The details of the experiments with the mouse control version of the DA Game are presented in our previous work [24]. For the purposes of this study we selected 10 consecutive trials from the intermediate phase to compare them in all of our analyses with the 10 gaze control trials. The difficulty of the game was set to a default value, meaning a decrease of 15% on average compared to the last trial selected for comparison.

The performance of the participants is determined by the user errors, which can be separated into two categories:

²<https://tobiigaming.com/product/tobii-eyex/>

- (i) *errors of omission* are the cases when the player misses an action; these are the more common ones and can have several causes such as the place of the action is outside the visual field or too little time to handle multiple parallel tasks;
- (ii) *errors of commission* occur when the player performs a wrong action, and does not correct it. These mistakes are the more rare ones in the DA Game and can happen when the player confuses two colors, performs an action too early or acts recklessly because of pressure.

In our analysis we computed the number of each type of user error and compared the means between the mouse control and the gaze-based control versions using the repeated measures analysis of variance (ANOVA) statistical model. Furthermore, we fit linear regression lines on the number of user errors to analyze the change over the trials from our experiments.

We calculate the length of the time intervals passed from the moment of a proper switch flip until the square actually passes the switch node. This latter event corresponds to the last moment when the switch could have been still flipped. We compared the distributions of these remaining time intervals between the mouse and the gaze-based control versions, to see whether there are considerable differences, i.e. whether the dwell time is limiting performance.

We also analyze the strategies of the participants. Particularly, we define two measures that characterize their decision making. The first one, called *double switch*, refers to flipping the same switch twice in a row, where after the first proper switch the player fails to look away and the dwell time elapses again resulting in another erroneous switch flip. This action corresponds to performing a double click with the mouse. We fit linear regression lines on the number of double switches to check whether they show an increasing or decreasing pattern during the trials of our experiments.

The second strategic measure is called *planning* or *planning ahead* and is defined in detail in our previous work [24], where it was found the most important predictor of performance in a regression analysis. It involves thinking in advance, executing an action before the situation would become critical and has the effect of reducing future timing constraints and/or cognitive load. We compare the planning strategic measure between the mouse and gaze-based control experiments using repeated measures ANOVA.

In our analysis we test the following experimental hypotheses.

H1 Despite carefully controlling experimental and design aspects, the number of user errors is considerably increased compared to the mouse control version.

However, we expect to observe a slow decrease in the number of errors and in the number of double switches over time in the amount of experiments we performed, i.e. players would start to get used to the gaze-based control version of the DA Game.

H2 The 500 ms dwell time does not influence considerably the distribution of the

remaining times, i.e. it does not impair performance by limiting the available times to perform switch flips.

H3 The planning strategic measure is decreased when using gaze-based control, most likely because of the higher cognitive load.

3 Results

We calculated the total number of user errors for each trial in both conditions (mouse control and gaze control). There was a statistically significant difference in means of user errors between the two control types, as determined by the repeated measures ANOVA, $F(1, 8) = 61.19$, $p < 0.001$. Figure 2 shows the mean of the user error numbers across trials in each of the two control versions of the DA Game, separately for every subject. Clearly, gaze control yielded lower scores on average.

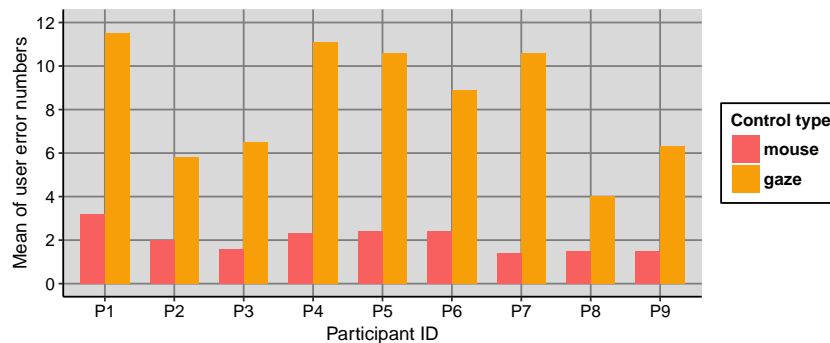


Figure 2: Comparison of overall mean of user error numbers, separately for participants.

The proportion of the commission type of errors to the total number of errors was also calculated in both conditions. Figure 3 compares these percentages computed over all 10 trials for each participant. We can see that generally the proportion of commission errors is considerably higher in the gaze version. Also, the repeated measures ANOVA for proportion of commission errors (computed for each trial separately) showed significant main effects of control type (mouse vs. gaze), $F(1, 8) = 20.28$, $p = 0.002$.

Figure 4 shows regression lines fitted on the number of user errors across trials. The errors of omission are decreasing in case of six subjects (P1, P3, P5, P6, P7, P9) and increasing in case of two subjects (P2 and P8). The errors of commission are decreasing in case of five (P3, P5, P7, P8, P9) and increasing in case of two subjects (P4 and P6). For some players we can see a reasonable learning rate when considering the sum of errors (P3, P5, P7, P9). However, the average of the

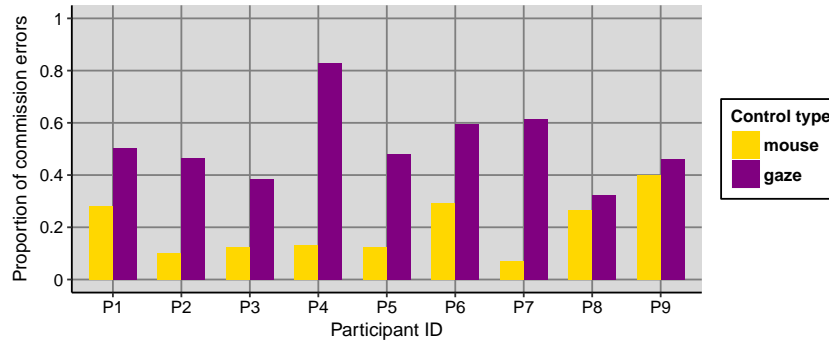


Figure 3: Comparison of overall proportion of commission type user errors, separately for participants.

user error numbers still remained considerably higher when compared to the mouse control version, as seen on Figure 2.

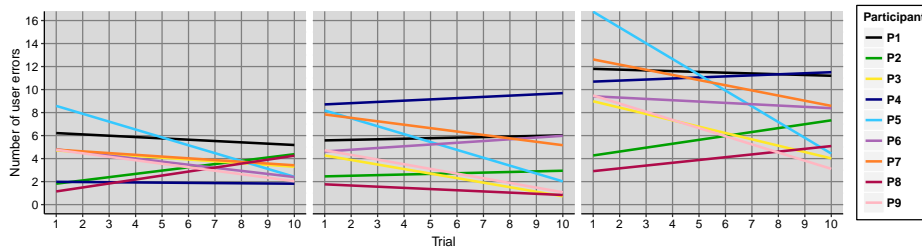


Figure 4: Linear regression lines fitted on the number of errors separately for participants, from left to right: errors of omission, errors of commission, sum of the two error types.

Figure 5 visualizes the smoothed distributions of the remaining time interval lengths, for both the mouse and gaze versions. We used the two-sample one-sided Mann-Whitney-Wilcoxon test to assess whether there are statistically significant differences between the distributions of the data for each of the 9 participants. The tests indicated that the remaining times were greater for the mouse control version than for the gaze-based control version in case of only 3 participants (P1, P8 and P9), $p < 0.001$. Inspecting the graphs on Figure 5 we can see that the distributions are surprisingly similar in case of each participant, the peaks for mouse control are higher and the curves are not shifted consistently to the left when switching to gaze-based control.

Figure 6 shows regression lines fitted on the number of double switches across trials. The lines show a considerable decrease for P5 and P7, and a substantial increase for P3.

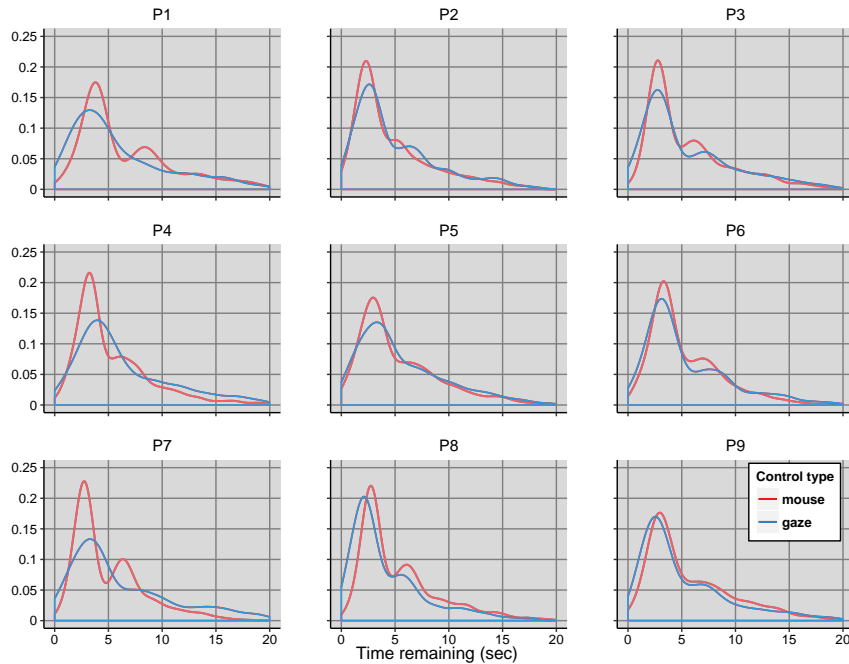


Figure 5: Comparison of distributions of remaining time intervals from the moment of switch flip until last possibility of performing the action, separately for participants. The title of each subplot is the identifier of the participant.

Figure 7 shows the comparison of the planning strategic measure between the mouse and gaze control versions for every participant. The results are mixed, confirmed also by repeated measures ANOVA, which revealed no significant main effect of control type, $F(1, 8) = 1.35$, $p = 0.28$. Particularly, planning was increased in case of four subjects (P4, P5, P6, P7), while it decreased for the others when switching to gaze-based control.

4 Discussion

We investigated the effect of switching from mouse control to gaze-based control in a complex divided attention task. Three hypotheses were tested and we elaborate the findings below.

H1 This was confirmed, since the number of user errors was increased in the gaze control version as shown on Figure 2, and this change was also statistically significant, as determined by the repeated measures ANOVA.

Regarding the second part of this hypothesis, it was confirmed only partially. Not all participants started to get used to the gaze-based control version of the

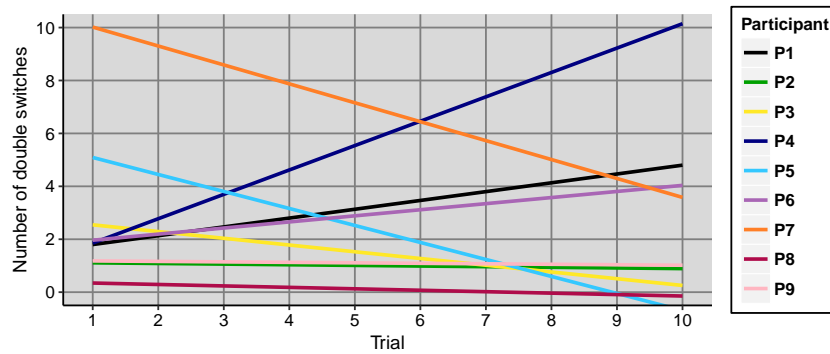


Figure 6: Linear regression lines fitted on the number double switches separately for participants.

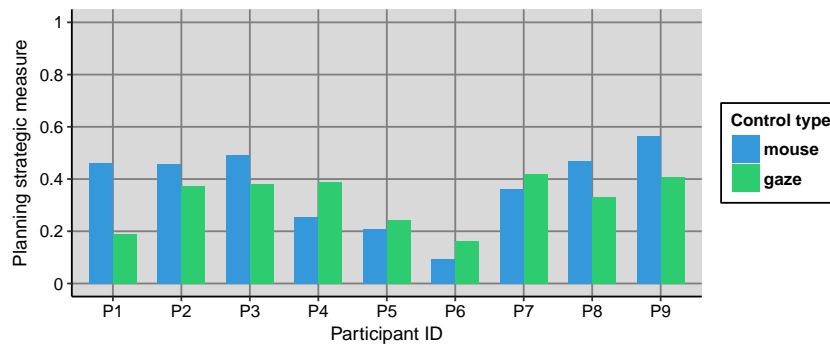


Figure 7: Comparison of planning strategic measure separately for participants.

DA Game, because the total number of user errors and the number of double switches decreased during our experiments only in the case of P3, P5, P7 and P9, as demonstrated by Figures 4 and 6.

H2 This was confirmed, i.e. the 500 ms dwell time did not limit considerably the available times to perform switch flips. Although the Mann-Whitney-Wilcoxon test indicated statistically significant differences between the distributions of the remaining times for the mouse and gaze control versions for three subjects, the distributions were similar for each participant, as shown on Figure 5, and the differences were not consistently reflecting the dwell time.

H3 This was confirmed only partially, because the planning strategic measure was actually increased in case of four participants, as seen on Figure 7.

Several notes have to be made regarding our experiments and results. The sum of the user errors was considerably increased in the gaze-based control version of

the DA Game. But what was really unanticipated is the significant increase of the proportion of errors of commission, as seen on Figure 3. This demonstrates that performance decrements can be attributed in large part to reckless actions from the increased cognitive load and not time constraints. Our subjects had difficulty in restraining themselves to not double-check the switches nodes after they made the proper action. Also they fail to avoid double switches, although the dwell time of 500 ms should be long enough to react and look away after the switch has been flipped once [29].

One might argue that our results are biased because of the selection of the trials for comparison from the mouse control version. However, the intermediate phase was the part where subjects were familiar with the DA Game and could play comfortably after the beginner phase [24]. In the advanced phase difficulty was high in order to test the effects of time pressure, so this would not provide a proper comparison. In addition, the difficulty of the DA Game was decreased in the experiments presented here.

The sample size in this study is small, which restricts the strength of generalizability of our findings and the statistical power of our analysis. It is plausible that the performance differences are due to the lack of practice with gaze-based control. A balanced study, where subjects would get experienced in gaze-based control first, is almost impossible to perform. Nevertheless, analyzing the strategies of participants shows that it is possible to achieve a fast learning rate in the gaze version of the DA Game. Specifically, the number of both user error types and the number of double switches were decreasing during our experiments and also the planning strategic measure was increased compared to mouse control in case of subjects P5 and P7 (see Figures 4, 6 and 7). One important lesson to learn from our experiments is that since using gaze as an input method in dynamic environments requires conscious effort from the user to carefully avoid looking at prohibited targets, the implementation of such interfaces requires careful design and experimentation.

4.1 Future work possibilities

It can be argued that choosing dwell time as the selection method in gaze-based control can limit performance. Indeed this latency contributes to cognitive load because it limits the exploration time in the DA Game. The choice of dwell time duration represents a trade-off between speed and accuracy. To achieve better performance adaptive dwell time [5, 15, 16] might represent a plausible alternative. This may require machine learning techniques in order to find a good model for predicting dwell times in our dynamic task.

One can implement an animation to indicate the progression of dwell time [15, 18]. Also the item in focus might be highlighted to increase the responsiveness of the interface. However, these could distract attention in spatial tasks with timing constraints. Also zooming into the focus area might be counterproductive as it means losing context information too [13].

An alternative for dwell time based selection is using a blink for the signal. But this would disrupt the natural interaction by requiring the user to think about it

before blinking. Another option for performing a click is to combine gaze-pointing with a hardware button. This may be faster than simple dwell time, but less accurate since users may tend to click before gaze has fully settled on the target [14]. One could combine gaze with other input modalities [17], for example speech, head movements or even feet [3]. However, these would not work for people with disabilities who could potentially use only gaze as an input method.

Gaze gestures [11, 17, 26] might provide a robust alternative to dwell-based interaction to avoid unintentional commands. Some researchers also studied pre-defined gaze patterns as pre-programmed strategies for control [7, 14]. However, participants with disabilities would have considerable difficulty in performing such gaze patterns or gestures.

Gaze input requires concentration to control the eyes consciously. Implicit use of gaze for control can release users from this burden [3], while explicit gaze input should be applied carefully, since it may have cognitive drawbacks. Combining gaze-based control with EEG signal analysis can help to gain further insight into cognitive processes [25].

All these alternative options for employing gaze-based control require further studies, possibly using a larger sample size.

5 Conclusion

In this paper, we conducted a small scale experimental study to analyze the impact of switching from mouse to gaze-based control in a special divided attention task, requiring continuous focused concentration and frequent shifts of attention. We conducted experiments with 9 participants and carefully controlled design and experimental aspects: the mouse control version of the task was well practiced, the difficulty was adjusted to the more demanding conditions and the parameters of gaze-based control were selected based on previous research findings. Despite all these circumstances, gaze control had a significant negative impact on the performance of participants.

In contrast to our assumptions, experienced users could not get used to gaze-based control in the amount of experiments we performed. On the other hand, by investigating the problem solving strategies of users, we showed that some subjects could make considerable progress in our task even in a short amount of practice. Our efforts suggest that with careful design, proper testing and sufficient user training, gaze controlled computer interfaces can become helpful in environments requiring divided attention.

Acknowledgements

The authors would like to thank their supervisor, András Lőrincz, for his expertise during the project, Zoltán Tóser for supervising the implementation process of the divided attention game used in this study, and Richárd Bellon and Dávid Hornyák

for implementing the software. The authors also express their gratitude to the volunteers who kindly participated in the experiments.

References

- [1] Bednarik, Roman, Gowases, Tersia, and Tukiainen, Markku. Gaze interaction enhances problem solving: Effects of dwell-time based, gaze-augmented, and mouse interaction on problem-solving strategies and user experience. *Journal of Eye Movement Research*, 3(1), 2009.
- [2] Borji, Ali, Lennartz, Andreas, and Pomplun, Marc. What do eyes reveal about the mind? algorithmic inference of search targets from fixations. *Neurocomputing*, 149:788–799, 2015.
- [3] Çöltekin, Arzu, Hempel, J., Brychtova, A., Giannopoulos, Ioannis, Stellmach, Sophie, and Dachsel, Raimund. Gaze and feet as additional input modalities for interacting with geospatial interfaces. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume III-2. ETH-Zrich, 2016.
- [4] Chen, Chun-Ching and Huang, Yen-Yi. Exploring the effect of color on the gaze input interface. In *2018 IEEE International Conference on Applied System Invention (ICASI)*, pages 620–623, April 2018.
- [5] Chen, Zhaokang and Shi, Bertram E. Using variable dwell time to accelerate gaze-based web browsing with two-step selection. *International Journal of Human-Computer Interaction*, 2018.
- [6] Dorr, Michael, Pomarjansch, Laura, and Barth, Erhardt. Gaze beats mouse: A case study on a gaze-controlled breakout. *PsychNology*, 7(2):197–211, 2009.
- [7] Fedorova, Anastasia A., Shishkin, Sergei L., Nuzhdin, Yu O., and Velichkovsky, Boris M. Gaze based robot control: The communicative approach. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 751–754, April 2015.
- [8] Gibaldi, Agostino, Vanegas, Mauricio, Bex, Peter J., and Maiello, Guido. Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behavior Research Methods*, 49(3):923–946, 2017.
- [9] Hardy, Joseph L., Farzin, Faraz, and Scanlon, Michael. The science behind Lumosity, Version 2, 2013. Lumos Labs, Inc.
- [10] Holmqvist, Eva, Derbring, Sandra, and Wallin, Sofia. Participation through gaze controlled computer for children with severe multiple disabilities. *Studies in Health Technology and Informatics*, 242:1103–1108, 2017.

- [11] Hyrskykari, Aulikki, Istance, Howell, and Vickers, Stephen. Gaze gestures or dwell-based interaction? In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 229–232, New York, NY, USA, 2012. ACM.
- [12] Jacob, Robert J. K. What you look at is what you get: Eye movement-based interaction techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, pages 11–18, New York, NY, USA, 1990. ACM.
- [13] Kern, Dagmar, Marshall, Paul, and Schmidt, Albrecht. Gazemarks: Gaze-based visual placeholders to ease attention switching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2093–2102, New York, NY, USA, 2010. ACM.
- [14] Lutteroth, Christof, Penkar, Moiz, and Weber, Gerald. Gaze vs. mouse: A fast and accurate gaze-only click alternative. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pages 385–394, New York, NY, USA, 2015. ACM.
- [15] Majaranta, Päivi, Ahola, Ulla-Kaija, and Špakov, Oleg. Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 357–360, New York, NY, USA, 2009. ACM.
- [16] Majaranta, Päivi, Aula, Anne, and Rähä, Kari-Jouko. Effects of feedback on eye typing with a short dwell time. In *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications*, ETRA '04, pages 139–146, New York, NY, USA, 2004. ACM.
- [17] Majaranta, Päivi and Bulling, Andreas. *Eye Tracking and Eye-Based Human-Computer Interaction*, pages 39–65. Springer, London, 2014.
- [18] Majaranta, Päivi, Isokoski, Poika, Rantala, Jussi, Špakov, Oleg, Akkil, Deepak, Kangas, Jari, and Raisamo, Roope. Haptic feedback in eye typing. *Journal of Eye Movement Research*, 9(1), 2016.
- [19] Majaranta, Päivi, MacKenzie, I. Scott, Aula, Anne, and Rähä, Kari-Jouko. Auditory and visual feedback during eye typing. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pages 766–767, New York, NY, USA, 2003. ACM.
- [20] Menges, Raphael, Kumar, Chandan, Sengupta, Korok, and Staab, Steffen. eyegui: A novel framework for eye-controlled user interfaces. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, NordiCHI '16, pages 121:1–121:6, New York, NY, USA, 2016. ACM.
- [21] Murata, Atsuo. Eye-gaze input versus mouse: Cursor control as a function of age. *International Journal of Human-Computer Interaction*, 21(1):1–14, 2006.

- [22] Park, Seonwook, Spurr, Adrian, and Hilliges, Otmar. Deep pictorial gaze estimation. In *The European Conference on Computer Vision (ECCV)*, pages 721–738, September 2018.
- [23] Prabhakar, Gowdham and Biswas, Pradipta. Eye gaze controlled projected display in automotive and military aviation environments. *Multimodal Technologies and Interaction*, 2(1), 2018.
- [24] Rill, Róbert Adrian, Faragó, Kinga Bettina, and Lőrincz, András. Strategic predictors of performance in a divided attention task. *PLOS ONE*, 13(4):1–27, 2018.
- [25] Sengupta, Korok, Sun, Jun, Menges, Raphael, Kumar, Chandan, and Staab, Steffen. Analyzing the impact of cognitive load in evaluating gaze-based typing. In *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 787–792, June 2017.
- [26] Shao, Yuan-Fu, Wang, Chiuan, and Fuh, Chiou-Shann. Eyelasso: Real-world object selection using gaze-based gestures. In *28th IPPR Conference on Computer Vision, Graphics, and Image Processing*, 2015.
- [27] Smith, J. David and Graham, T. C. Nicholas. Use of eye movements for video game control. In *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, ACE '06*, New York, NY, USA, 2006. ACM.
- [28] Stampe, Dave M. and Reingold, Eyal M. Selection by looking: A novel computer interface and its application to psychological research. In Findlay, John M., Walker, Robin, and Kentridge, Robert W., editors, *Eye Movement Research*, volume 6 of *Studies in Visual Information Processing*, pages 467–478. North-Holland, 1995.
- [29] Woods, David L., Wyma, John M., Yund, E. William, Herron, Timothy J., and Reed, Bruce. Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience*, 9:131, 2015.
- [30] Yarbus, Alfred L. *Eye movements and vision*. Plenum Press, 1967.