



Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor

Shaoyao Zhang, Yu Tian, Chunhui Wang & Kunlin Wei

To cite this article: Shaoyao Zhang, Yu Tian, Chunhui Wang & Kunlin Wei (2020): Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor, Ergonomics, DOI: [10.1080/00140139.2020.1762934](https://doi.org/10.1080/00140139.2020.1762934)

To link to this article: <https://doi.org/10.1080/00140139.2020.1762934>



Accepted author version posted online: 29 Apr 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor

Shaoyao Zhang^{a#}, Yu Tian^{ab#}, Chunhui Wang^a, Kunlin Wei^{bcd*}

^aNational Key Laboratory of Human Factors Engineering, China Astronaut Research and Training Center, Beijing, China

^bSchool of Psychological and Cognitive Sciences, Peking University, Beijing, China

^cBeijing Key Laboratory of Behavior and Mental Health, Beijing, China

^dKey Laboratory of Machine Perception, Ministry of Education, Beijing, China

^ePeking-Tsinghua Center for Life Sciences, Peking University, Beijing, China

[#]ShaoyaoZhang and Yu Tian contributed equally.

[[#]These authors contributed equally to this work.]

*Correspondence should be addressed to K.W. (wei.kunlin@pku.edu.cn)

Kunlin Wei, School of Psychological and Cognitive Sciences, Peking University, Beijing 100080, China.

Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor

Abstract: Eye movement-based human-computer interactions are emerging in diverse scenarios. When selecting targets on a user interface, the method of combining fast gaze pointing with reliable manual action is becoming increasingly popular. However, this method suffers from noise in gaze pointing caused by eye jitters and users' habitual early move-away of gaze before manual actions. Here we propose a novel solution to mitigate these problems by locking the gaze cursor at the target for imminent manual selection. We compared this gaze-lock cursor with a conventional gaze cursor in a typing task with varying key sizes and key gaps. Results show that typing performance was significantly better with larger key size and gap. More importantly, the gaze-lock cursor significantly increased speed and decreased errors when compared to a conventional gaze cursor. Our findings demonstrate that the gaze-lock cursor is a promising tool for gaze interactions involving frequent target selections.

Practitioner Summary: Target selection by gaze pointing and manual confirmation suffers from eye jitters and users' habitual early move-away of gaze before manual actions. The performance of this method can be improved by applying the gaze-lock cursor we proposed, increasing target size or increasing the target gap.

Keywords: eye tracking, human-computer interaction, gaze cursor, eye jitters

Introduction

Gaze tracking has been proposed as a promising tool for human-computer interactions (Duchowski 2018; Miyoshi and Murata 2001; Murata 2006; Wang, Zhai, and Su 2001; Zhai, Morimoto, and Ihde 1999). While early studies used gaze tracking techniques to facilitate interactions of motor disabled individuals with computers (Bates, Vickers, and Istance 2010; Chin et al. 2008; Pinheiro et al. 2011), recent advances in the field have extended gaze-based interactions to various scenarios and tasks, including virtual reality and mixed reality (Blattgerste, Renner, and Pfeiffer 2018; Choe et al. 2019), mobile devices (Bulling and Gellersen 2010), large screen displays (Cho et al. 2012), and robots (Yu et al. 2014).

Target selection is a common and basic operation in various kinds of human-computer interactions. It often refers to the process of moving towards a target (e.g., a certain key on a virtual keyboard) followed by an action selection (e.g., click a mouse button to input the letter on the key). Target selection by gaze pointing can be achieved by requiring the user to fixate on the target for a predetermined time interval, i.e. the dwell time. This dwell-based approach involves two components, i.e., a goal-directed saccade to the target and a period of sustained fixation. The advantage of the dwell-based approach is that it works by using only the eyes, without involvement of hands or feet. This is particularly important for individuals with specific motor disabilities. However, this method often leads to unintended selections as it does not distinguish between whether the user intends to select the fixated object or just looks at it for an arbitrary reason. This problem has been acknowledged for a long time and is referred to as the Midas-Touch problem (Chin et al. 2008; Jacob 1991). Setting a long dwell time may reduce the probability of unintended selections, but it will increase task completion time and visual workload (Jacob 1991; Mott et al. 2017).

To reduce the Midas-Touch problem and maintain efficiency, researchers came up with a common solution: use the gaze to point to the target and use another method for selection (Majaranta and Bulling 2014). The action selection could be achieved by manual clicking (Kumar et al. 2007; Murata and Karwowski 2019; Zhai, Morimoto, and Ihde 1999), foot pedaling (Klamka et al. 2015), head movement (Kurauchi et al. 2015), movement of facial muscles (Tuisku et al. 2012; Tuisku et al. 2013) or speech (Miniotas et al. 2006; Murata and Karwowski 2019). Among them, the combination of gaze pointing with manual clicking offers the best performance (Jochems, Vetter, and

Schlick 2013; Murata and Karwowski 2019), owing to speed and accuracy advantages of hand movements. However, the eye-hand combination input method also faces usability problems. For instance, the gaze point of the user is sometimes not right on the target at the time of manually clicking, resulting in a failure of target selection or even a wrong selection (Kumar et al. 2008). This usability problem may be attributed to several factors, such as eye jitters, measurement noise of the eye tracking system, and the habitual tendency to move the gaze away before the manual click.

Eye jitters refer to the inherent jumpy movements of eyes during fixations. These movements are usually involuntary with small amplitudes (Jacob 1991). In gaze-based interactions, the eye tracker can detect the user's gaze point on the screen in real time, if a cursor is displayed at the user's gaze point, eye jitters can make the gaze cursor visually distracting (Majaranta and Bulling 2014; Zhang, Feng, and Zha 2011). For the eye-hand combination input method, eye jitters at the time of clicking might cause errors (Kumar et al. 2008; Stellmach and Dachsel 2012). Besides eye jitters, when using desktop applications with a mouse, users tend to shift their gaze from the target before clicking (Aoki, Hansen, and Itoh 2008; Liebling and Dumais 2014). This tendency could lead to errors (i.e., clicking occurring after the gaze has shifted away from the target) when using the approach combining gaze pointing with manual clicking (Kumar et al. 2008).

Researchers have already proposed several methods to stabilize the gaze cursor and to reduce errors due to eye jitters. The first method is based on filtering the noisy gaze points in real time and only displaying the smoothed data (De Gaudenzi and Porta 2013; Kumar et al. 2008; Zhang, Ren, and Zha 2008), thus substantially improving the stabilization of the gaze cursor. Other methods, named the "warping to target center" (WTC) method (Zhang, Ren, and Zha 2008) and "automatic lock of cursor movement" (ALCM) method (Murata and Karwowski 2019), are based on displaying the gaze cursor to the target center once the gaze point falls in the target area, and releasing the gaze cursor to the gaze point once human gaze jumps out of the target. It should be mentioned that although both the WTC and ALCM methods follow similar principles, they produce quite different results. In Zhang's study, the participants were instructed to hold the gaze cursor on the target for 7 seconds to activate it. The times that the gaze cursor entered and exited the target were used to evaluate the stability of the gaze cursor. Results showed that, compared to the conventional gaze cursor mode in which the cursor follows the gaze point in real time, the cursor stability was actually worsened

when the WTC method was applied (Zhang, Ren, and Zha 2008). This performance decrement was caused by the fact that the gaze point entered and exited the target during long-time fixations, leading to abrupt cursor jumps when the WTC method was imposed. In contrast, Murata and Karwowski combined gaze pointing with another input modality (a voice command or a key press), and used more realistic tasks (dragging an object to a target, and selecting a menu item) to test the effectiveness of the ALCM method. They found that the ALCM method significantly improved the performance of the two tasks tested (Murata and Karwowski 2019). While the ALCM method appears promising as an eye-gaze input system, it still presents some limitations. For instance, when there are many alternatives closely arranged on the interface and the task requires visual search, the gaze cursor will be frequently locked on non-targets during the visual search processes or the saccade processes, which can be visually distracting. What is more, while tasks that require fast sequential selections of targets, such as the gaze typing, are common for eye-gaze input system, the method of locking the gaze cursor has not been tested in this kind of task.

To deal with the usability problems in eye-hand combination input, we proposed a novel method called the gaze-lock method for locking and unlocking the gaze cursor at the target center with temporal constraints. The proposed gaze-lock method was compared to a conventional gaze cursor in a typing task with a virtual keyboard. We varied the key width and key gap to comprehensively evaluate its effectiveness in dealing with visual search on an interface with crowded potential targets. The results have implications for the design of interfaces for gaze interactions.

Methods

Participants

Thirty-two adults (9 females) were recruited. The participants' average age was 26.5 ± 3.5 (age range 22-37 years). Each participant received 70 Yuan (equivalent to about 10 dollars) for their participation. All participants were right-handed, with normal or corrected-to-normal vision, without any known history of psychiatric, movement or neurological disorders. All participants were experienced computer users with no prior experience of eye tracking. The study was approved by the ethics committee of China Astronaut Research and Training Center. All participants provided an informed written

consent prior to formal data collection.

Apparatus

The eye tracker (Tobii 4C, with a purchased research license) was attached on the bottom of a 23 inch LCD screen (DELL P2317H). The LCD screen was placed on a height adjustable desk (E1D, Loctek Ergonomic Technology Corp). Eye movement acquisition rate was 90 Hz. Participants controlled the cursor on the screen by gaze, and clicked the left key of the trackball (Logitech MX ERGO) to confirm the selection (as shown in Figure 1).



Figure 1. Test setting for the experiment

Gaze cursor modes

In the present study, two cursor modes were designed for cursor display: a conventional gaze cursor and a gaze-lock cursor. The conventional gaze cursor was displayed as a blue hollow circle with a diameter of about 0.9 cm and updated its position based on the gaze position (Figure 2A-C). Since we applied the Exponential Weighted Moving Average (EWMA) algorithm (Stawicki et al. 2017) to smooth the original gaze point data, the cursor position update interval was 33ms and not of 11ms (the sample interval of the eye tracker), allowing for a 3 samples time window (33ms).

In the gaze-lock cursor mode, two modifications were introduced to the conventional gaze cursor. First, a temporal constraint for locking the gaze cursor was imposed: when the gaze cursor stayed in the target area beyond a predefined period of time (T_{lock}), it becomes locked on the target center. Once the cursor was locked, the

cursor remains on the target center, as long as the gaze point remained in the target area (Figure 2D & E). In this way, the visual effect of eye jitters which is frequently observed with the conventional gaze cursor could be reduced (Figure 2B). Second, another temporal constraint for unlocking the gaze cursor was imposed: when the gaze point fell out of the target area, the locked cursor would not be unlocked unless the gaze point left the target beyond a predefined period of time (T_{unlock}) (Figure 2F). If the gaze point returned to the target area within T_{unlock} , the locked cursor was not affected and still stays on the target center. This temporal constraint could handle the case that the jittery gaze point left the target area temporarily and unintentionally. Based on our pilot tests, T_{lock} and T_{unlock} were both set to 100ms.

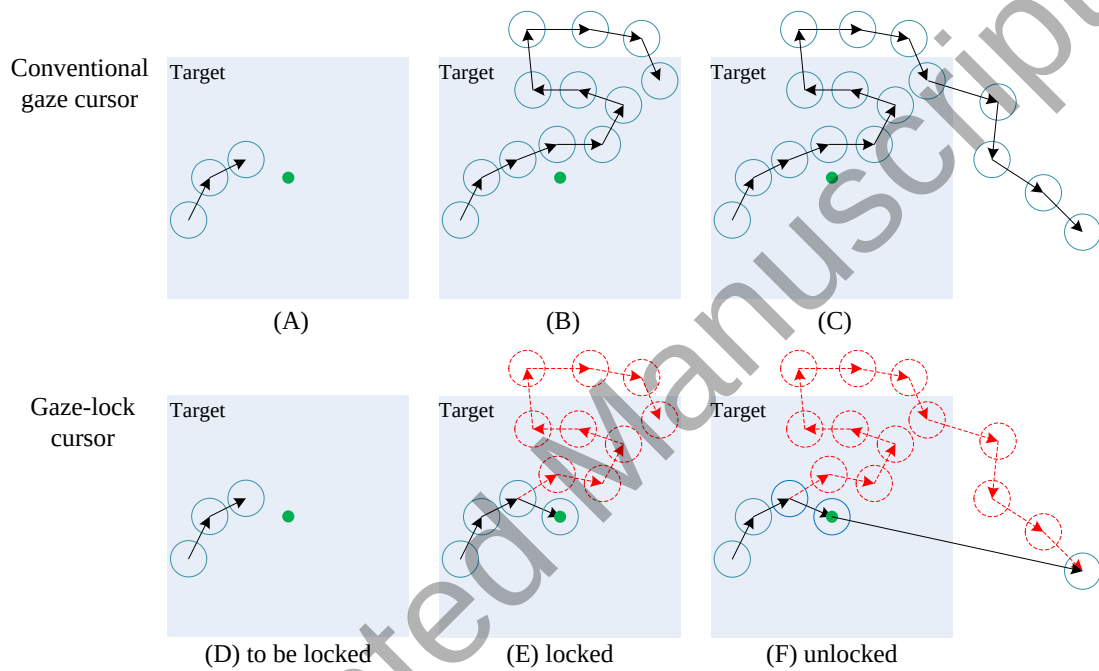


Figure 2. Illustrations of the two gaze cursor modes when displaying the same exemplary sequence of gaze points. The light blue rectangle denotes the target, and the green dot indicates the target center. The solid blue circles denote the displayed cursor positions, while the dashed red circles (appearing in E and F) denote the gaze positions that are disassociated with the displayed cursor position (locked at the target center). Solid lines with arrows show the trajectory of the displayed cursor, while dashed lines with arrows (appearing in E and F) show the trajectory of the gaze point. Note that the hollow circle, no matter solid or dashed, is based on the gaze point data smoothed by the EWMA algorithm, and updated every 33ms in this study. A-C) In the conventional cursor mode, the displayed cursor always follows the gaze point. In this exemplary trial, the participant tried to fixate on the target, but the gaze point wandered around the

target region due to eye jitters. D-F) The gaze-lock cursor mode. D) When the gaze point entered a target for less than T_{lock} (100ms), the cursor followed the gaze point. E) Once the gaze stayed on the target area over T_{lock} , the cursor (solid circle) was locked on the target center, irrespective of the actual gaze point (dashed red circle), as long as the gaze point did not leave the target over T_{unlock} (100ms). F) When the gaze point left the target over T_{unlock} (100ms), the cursor was unlocked and started following the gaze point again.

Experimental Task

Our task consisted of selecting a highlighted character by manually clicking the mouse while the gaze cursor was on the character key (Figure 3). This task was similar to the guided typing task used by Tula and colleagues (Tula, Kurauchi, and Morimoto 2013). The cursor was displayed as a blue hollow circle whose position was controlled by gaze. If the character was correctly selected within 3s, a beep sound was played and the color of the selected key was simultaneously changed to blue for about 200ms. Once the character was correctly selected or when the mandatory input time of 3s elapsed, whichever came first, the next character was immediately highlighted. The task terminated when either of the two conditions was satisfied: (1) the 36 target characters had all been highlighted in sequence; (2) the total time exceeded 45s. The 36 characters included 26 letters (from A to Z) and 10 numbers (from 0 to 9), and their order of presentation was randomized. For this modified guided typing task, we used sequences of randomly presented characters to minimize the practice effect and the influence of typing proficiency among individuals.

Square keys were used, and the levels of key width and key gap were selected (see below) according to the normal range of key sizes and key gaps commonly used in software such as Microsoft Word, Adobe Acrobat, and the QWERTY virtual keyboard provided on Microsoft Surface tablets.



Figure 3. A screen shot of the task interface

Experimental procedures

Our experiment used a $2 \times 2 \times 2$ within-subject repeated measures design. The first factor was the gaze cursor mode (the gaze-lock cursor vs. the conventional gaze cursor). The second and third factors were key width (1.2cm vs. 2.2cm) and key gap (0.2cm vs. 1.2cm). Thus, each participant was tested for a total of eight conditions. We selected 8 random sequences of characters with the same total travel distances, one sequence for each condition. The total travel distance for each sequence was exactly 120 unit of key-to-key distance. One unit of key-to-key distance is defined as the displacement from one key to its most adjacent key.

Participants were seated about 58 cm from the display, thus the width of 1cm on the screen corresponded to about 1 degree of visual angle (1°) in the present study. Participants were seated without a chin-rest and were instructed to adjust the height of the desk thus allowing them to look horizontally at the center of the vertical display (Figure 1). For each participant, the experimenter briefly introduced the eye tracking technique followed by an explanation of the task. After performing a 7-point calibration of the eye tracker, the participant began a practice session, and then carried out a formal test session.

Each participant was required to complete eight sequences corresponding to the eight task conditions. The order of the eight conditions was counterbalanced across participants using a Latin-square design. The time limit for each sequence was 45s in the formal test session. Participants were allowed to rest between conditions during the formal test. The entire experiment lasted approximately 40 minutes.

Performance Measures

To assess the performance across conditions, we defined four measures for general typing performance: error rate, execution time, edit distance, and overall input speed. We also analyzed the time spent on two distinct phases (the gaze shift phase and the click phase, see below) of the eye-hand combination input process to assess whether there was a potential advantage of the gaze-lock cursor over the conventional gaze cursor.

Error rate (ER) was defined as the ratio of error selection times to total selection times. If the selected character was not the highlighted one, it was recorded as an error selection. Lower error rate indicated better performance.

Execution time (ET) was defined as the completion time of the whole task. The longest possible ET was 45s as pre-defined by our task setting. Shorter ET indicated better performance.

Edit distance (ED) was defined as the minimum number of operations required to transform the typed string, possibly with errors (such as wrong input, missed input, or extra input), into the predefined target string (He et al. 2014; Meister 2015; Tian et al. 2018). For example, if the predefined target string was “YES”, and the actual string typed in was “YDS” or “YS” or “YESS”, the ED would be 1 in all three cases. If the actual string was exactly the same as the target string, then the ED would be 0. Smaller ED meant fewer errors and better performance.

Using ED as an approximation of the number of errors, we computed the number of characters that were correctly typed within a minute. This index was referred to as input speed, which was computed as $60 \times (36 - ED) / ET$. Note, 36 was the total number of characters in the target string, thus $(36 - ED)$ denoted the number of characters that were correctly typed. Input speed was thus a comprehensive performance measure which combined ED and ET. Higher input speed indicated better performance.

We divided the process of inputting a character into two successive phases, the gaze shift phase and the click phase. The gaze shift phase was defined as the time from the moment in which the desired key was highlighted up until the displayed gaze cursor first landed on the highlighted key. The click phase was defined as the time from the moment in which the cursor first landed on the key to the moment in which a manual click was performed. The duration of the two phases was defined as gaze shift time

(T_{shift}) and click latency (T_{click}), respectively. We were interested not only in the mean duration of two phases, but also their variances across characters. The character-to-character variance could reflect how fluent participants were able to input the target string as it represents fluctuations of time spent on successive gaze shifts and manual clicks. We computed coefficients of variation (CV), instead of standard deviations, of T_{shift} and T_{click} across characters to normalize individual differences. Lower CV was assumed to denote more fluent performance.

Results

Overall, the gaze-lock cursor led to significantly better typing performance. Larger key width and key gap also led to better typing performance. We report the results for each dependent measure below.

Input speed

The input speed was significantly higher with the gaze-lock cursor compared to the conventional gaze cursor, as the number of input characters per minute increased from 53.0 ± 2.2 to 59.0 ± 1.7 ($p = 0.000$, an average improvement of 11.2%). Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 33.820$, $p = 0.000$, $\eta^2 = 0.522$), key width ($F(1, 31) = 86.291$, $p = 0.000$, $\eta^2 = 0.736$) and key gap ($F(1, 31) = 98.253$, $p = 0.000$, $\eta^2 = 0.760$). The input speed was significantly higher under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap.

We found a significant interaction between gaze cursor mode and key width ($F(1, 31) = 5.313$, $p = 0.028$, $\eta^2 = 0.146$), and between key gap and key width ($F(1, 31) = 32.484$, $p = 0.000$, $\eta^2 = 0.512$). While all simple main effects were significant ($p = 0.000$), the interaction effects suggested that the effect of gaze cursor mode and the effect of key gap were larger if the key width was set on the relatively smaller level (1.2°). No other interaction was significant.

As shown in Figure 4, the input speed increased as the key width and the key gap increased. The smaller key width (1.2°) with the smaller key gap (0.2°) led to the slowest speed, while the larger key width (2.2°) with the larger key gap (1.2°) led to the fastest speed. Meanwhile, the advantage brought by the gaze-lock cursor was significant

in three of the four conditions except for the easiest condition with the larger key width (2.2°) and the larger key gap (1.2°).

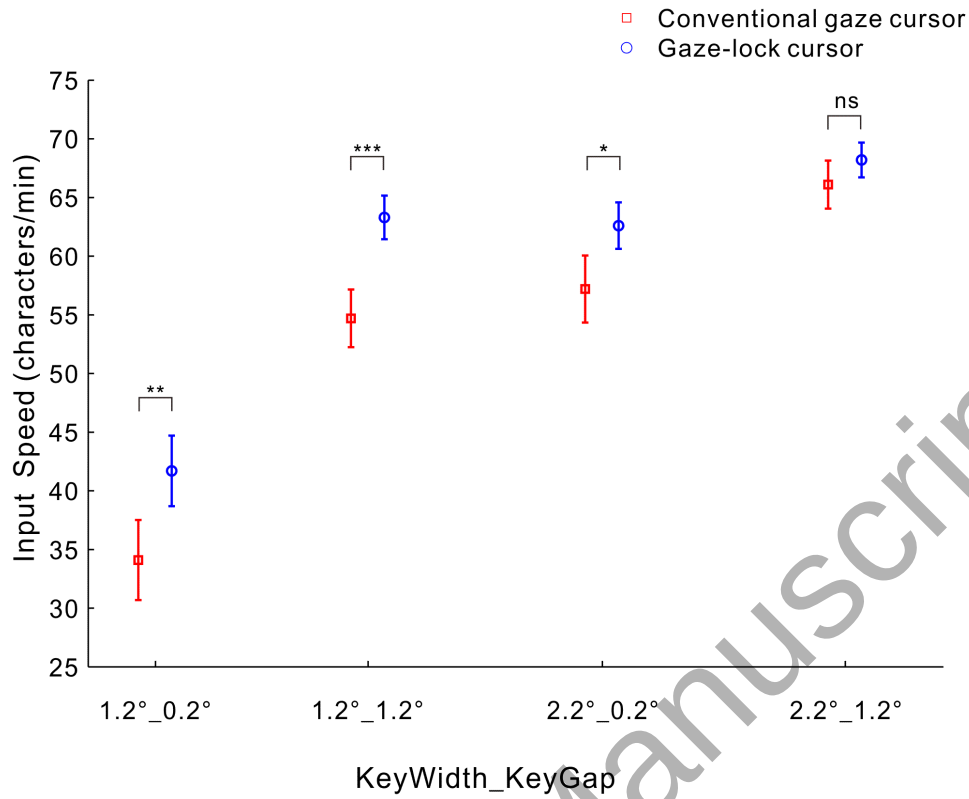


Figure 4. Average input speed under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Error rate

The error rate was significantly lower with the gaze-lock cursor compared to the conventional gaze cursor, decreasing from 0.104 ± 0.012 to 0.057 ± 0.007 (an average decrease of 45.0%). This indicates that participants made fewer selection errors with the gaze-lock cursor. As the error rate data did not conform to a normal distribution, the ART nonparametric test was applied (Wobbrock et al. 2011), and significant main effects of gaze cursor mode ($F(1, 217) = 48.657, p < 0.001$), key width ($F(1, 217) = 102.163, p < 0.001$) and key gap ($F(1, 217) = 142.874, p < 0.001$) were found. The error rate was significantly lower under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap.

We found significant two-way interactions between gaze cursor mode and key width ($F(1, 217) = 14.576, p < 0.001$), between gaze cursor mode and key gap ($F(1, 217)$

= 21.944, $p < 0.001$), and between key gap and key width ($F(1, 217) = 58.917$, $p < 0.001$). The interaction effects indicated that the advantage of the gaze-lock cursor was larger with the smaller key width (1.2°) and the smaller key gap (0.2°).

As shown in Figure 5, the error rate decreased as the key width and the key gap increased. The smaller key width (1.2°) with the smaller key gap (0.2°) led to the highest error rate, while the larger key width (2.2°) with the larger key gap (1.2°) led to the lowest error rate. The advantage brought by the gaze-lock cursor was significant in three of the four conditions except for the easiest condition with the larger key width (2.2°) and the larger key gap (1.2°).

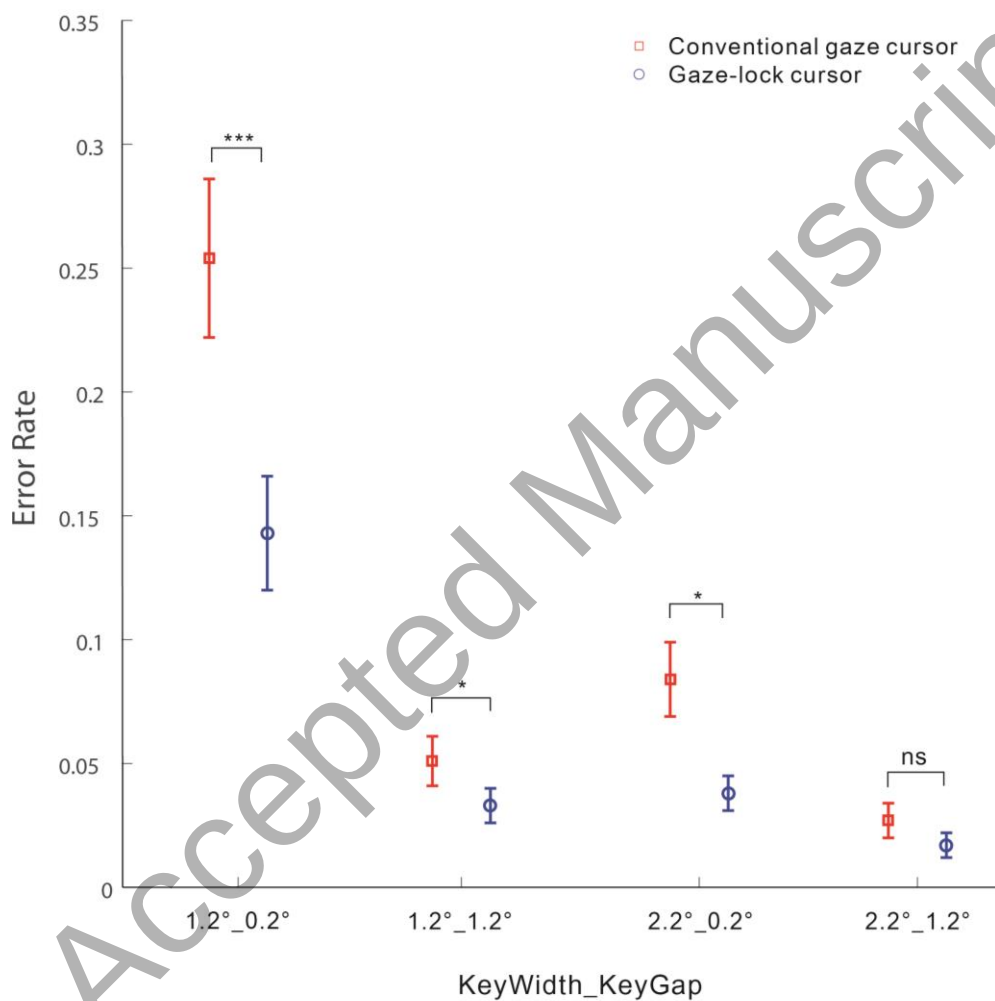


Figure 5. Average error rate under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Edit distance

The edit distance was significantly smaller with the gaze-lock cursor compared to the conventional gaze cursor, decreasing from 5.5 ± 3.1 to 3.1 ± 0.4 ($p = 0.000$, an average decrease of 44.0%). This indicated that the participants made fewer typing errors with the gaze-lock cursor. Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 25.514$, $p = 0.000$, $\eta^2 = 0.451$), key width ($F(1, 31) = 42.659$, $p = 0.000$, $\eta^2 = 0.579$) and key gap ($F(1, 31) = 57.078$, $p = 0.000$, $\eta^2 = 0.648$). The edit distance was significantly smaller under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap.

We found significant two-way interactions between gaze cursor mode and key width ($F(1, 31) = 5.601$, $p = 0.024$, $\eta^2 = 0.153$), between gaze cursor mode and key gap ($F(1, 31) = 7.827$, $p = 0.009$, $\eta^2 = 0.202$), and between key gap and key width ($F(1, 31) = 32.166$, $p = 0.000$, $\eta^2 = 0.509$). The interaction effects indicated that the advantage of the gaze-lock cursor was larger with the smaller key width (1.2°) and the smaller key gap (0.2°). The three-way interaction of gaze cursor mode, key width and key gap was not significant ($p > 0.1$).

As shown in Figure 6, the edit distance decreased as the key width and the key gap increased. The hardest condition with the smaller key width (1.2°) and the smaller key gap (0.2°) yielded the highest ED, while the easiest condition with the larger key width (2.2°) and the larger key gap (1.2°) yielded to the lowest ED. Meanwhile, the advantage brought by the gaze-lock cursor was significant in three of the four conditions except for the condition with the larger key width (2.2°) and the larger key gap (1.2°).

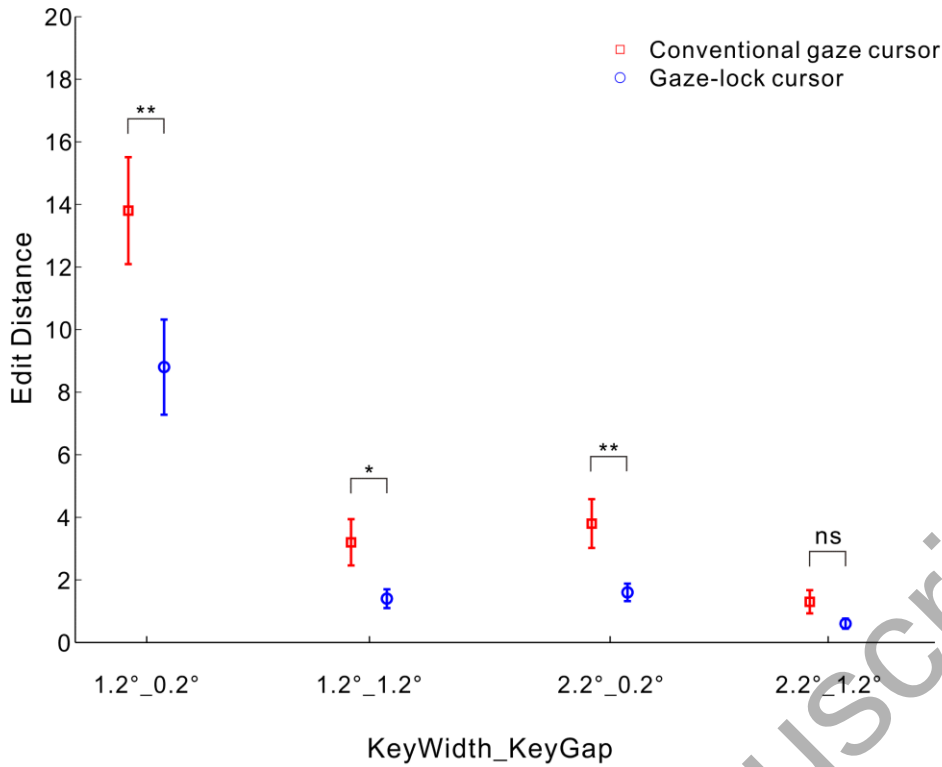


Figure 6. Average edit distance under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Execution time

The execution time was significantly shorter with the gaze-lock cursor compared to the conventional gaze cursor, decreasing from $36.8 \pm 0.9s$ to $35.0 \pm 0.7s$ ($p = 0.000$, an average decrease of 5.1%). Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 21.972$, $p = 0.000$, $\eta^2 = 0.415$), key width ($F(1, 31) = 73.795$, $p = 0.000$, $\eta^2 = 0.704$) and key gap ($F(1, 31) = 72.607$, $p = 0.000$, $\eta^2 = 0.701$). The execution time was significantly shorter under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap.

We found a significant two-way interaction between key gap and key width ($F(1, 31) = 11.592$, $p = 0.002$, $\eta^2 = 0.272$), and a significant three-way interaction of gaze cursor mode, key width and key gap ($F(1, 31) = 7.979$, $p = 0.008$, $\eta^2 = 0.205$). No other interaction was significant. The time-saving effect of the gaze-lock cursor manifested only in the two intermediate conditions, with key width and gap combinations of 1.2° & 1.2° and 2.2° & 0.2° (Figure 7). The easiest condition (width 2.2° & gap 1.2°) and the

hardest condition (width 1.2° & gap 0.2°) did not show the effect of the gaze-lock cursor mode.

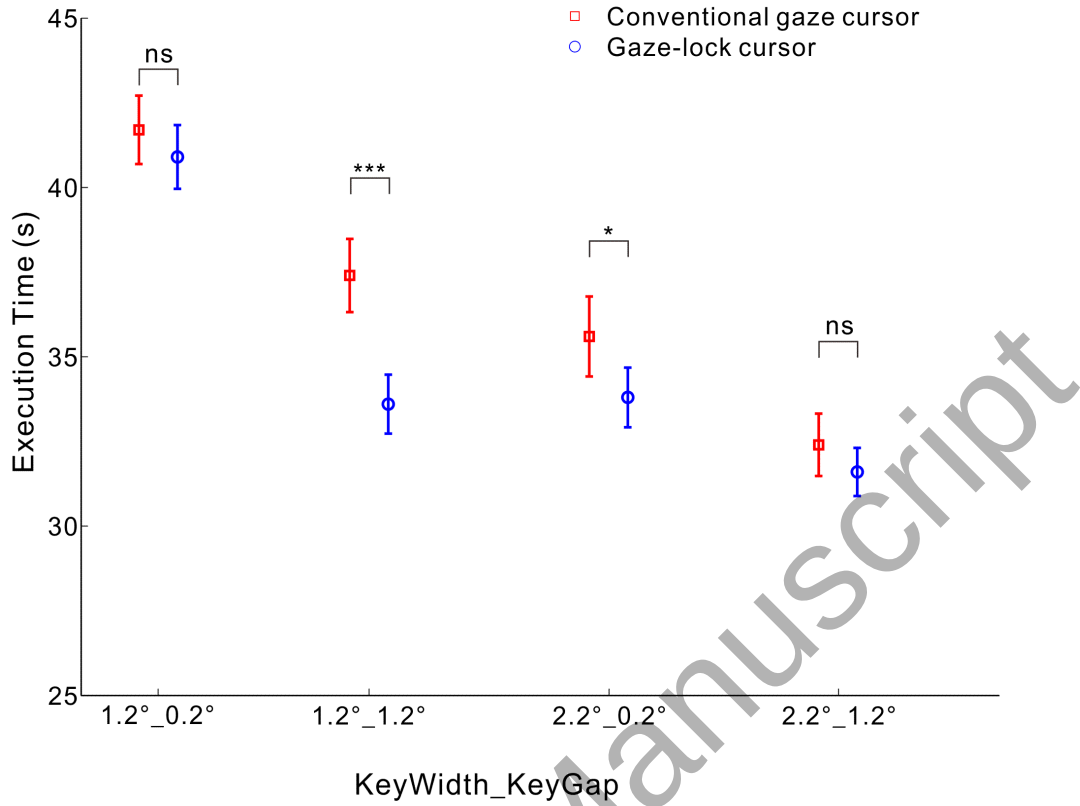


Figure 7. Average execution time under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Gaze shift time (T_{shift})

Our gaze-lock manipulation was only applied to the target selection phase. Therefore, we did not expect it to show any advantages over the conventional method in the duration of gaze shift. However, we found that T_{shift} was significantly shorter with the gaze-lock cursor than with the conventional gaze cursor, as T_{shift} decreased from $0.512 \pm 0.013\text{s}$ to $0.494 \pm 0.013\text{s}$ ($p = 0.000$, an average decrease of 3.5%). Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 10.993$, $p = 0.002$, $\eta^2 = 0.262$), key width ($F(1, 31) = 40.839$, $p = 0.000$, $\eta^2 = 0.568$) and key gap ($F(1, 31) = 60.540$, $p = 0.000$, $\eta^2 = 0.661$). T_{shift} was significantly shorter under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap (Figure 7). We found a significant two-way interaction between the key gap and key width ($F(1, 31) = 28.358$, $p = 0.000$, $\eta^2 = 0.478$). No other interaction was

significant.

We then analyzed the CV of T_{shift} across clicks. Three-way repeated-measures ANOVAs found significant main effects of key width ($F(1, 31) = 20.358, p = 0.000, \eta^2 = 0.396$) and key gap ($F(1, 31) = 87.098, p = 0.000, \eta^2 = 0.738$), but no main effect of cursor mode (0.367 ± 0.015 and 0.381 ± 0.015 for gaze-lock cursor and conventional gaze cursor respectively, $p = 0.162$). The CV of T_{shift} was significantly lower under the conditions of a larger key width, and a larger key gap (Figure 8). We found a significant interaction between key gap and key width ($F(1, 31) = 18.521, p = 0.000, \eta^2 = 0.374$). No other interaction was significant.

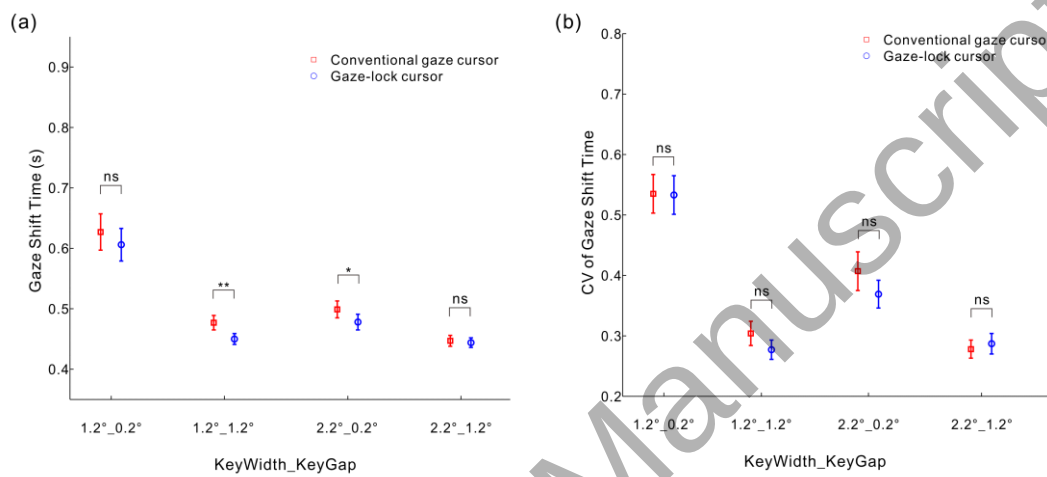


Figure 8. Average T_{shift} (a) and CV of T_{shift} (b) under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Click latency (T_{click})

Click latency (T_{click}) is the interval starting from the moment the gaze falls onto the target to the moment the click action is registered. T_{click} was significantly shorter with the gaze-lock cursor than with the conventional gaze cursor, decreasing from 0.606 ± 0.031 s to 0.522 ± 0.019 s ($p = 0.000$, an average decrease of 13.9%). Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 19.550, p = 0.000, \eta^2 = 0.387$), key width ($F(1, 31) = 67.841, p = 0.000, \eta^2 = 0.686$) and key gap ($F(1, 31) = 41.534, p = 0.000, \eta^2 = 0.573$). T_{click} was significantly shorter under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap (Figure 8). We found significant interactions between gaze cursor mode and key width ($F(1, 31) = 7.409, p = 0.011, \eta^2 = 0.193$), and between key gap and key width

($F(1, 31) = 23.810, p = 0.000, \eta^2 = 0.434$). The effect of gaze cursor mode and the effect of key gap were relatively larger with the smaller key width (1.2°).

We found that click latency showed significantly lower variance with the gaze-lock cursor than with the conventional gaze cursor. CV of T_{click} decreased from 0.590 ± 0.030 to 0.442 ± 0.028 ($p = 0.000$, an average decrease of 25.1%). Three-way repeated-measures ANOVAs found significant main effects of gaze cursor mode ($F(1, 31) = 42.343, p = 0.000, \eta^2 = 0.577$), key width ($F(1, 31) = 36.173, p = 0.000, \eta^2 = 0.539$) and key gap ($F(1, 31) = 57.019, p = 0.000, \eta^2 = 0.648$). The CV of T_{click} was significantly lower under the conditions of the gaze-lock cursor, the larger key width, and the larger key gap (Figure 9). No significant interaction was found ($p > 0.1$).

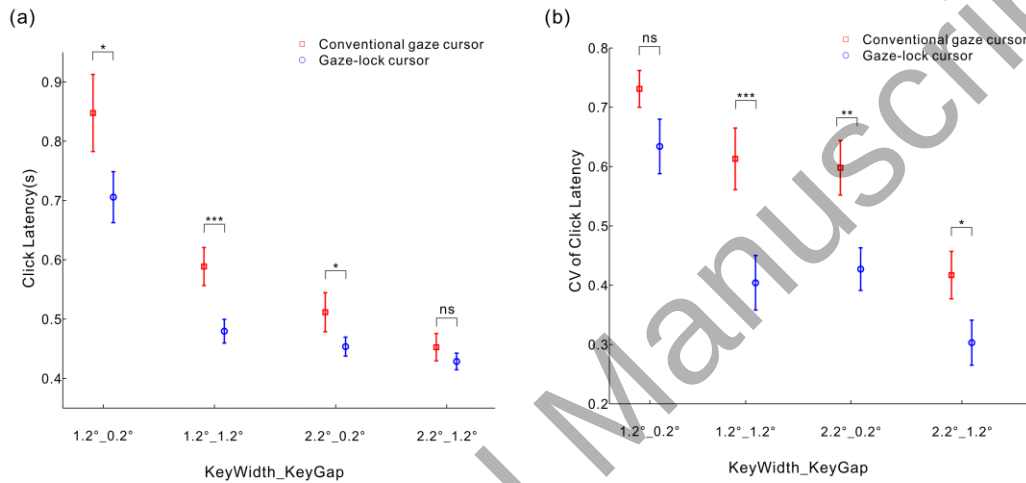


Figure 9. Average T_{click} (a) and CV of T_{click} (b) under different task conditions. Error bar denotes standard error. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant.

Discussions

In this study we compared the gaze-lock cursor with the commonly used conventional gaze cursor (Yu et al. 2017; Zhang, Feng, and Zha 2011; Zhang, Ren, and Zha 2008). We found that the input speed increased by 11.2% and the error rate decreased by 45.0% across conditions. Underneath this apparent performance boost, we found that most of the effect of the gaze lock cursor occurred during the manual selection of the target, i.e., the click phase. The duration from the moment the gaze cursor landed on the highlighted key to the moment this key was actually selected (T_{click}) was shortened dramatically by an average of 84ms per character. In contrast, the duration of the gaze shift phase (T_{shift}), from the moment the target key was highlighted to the moment the

gaze cursor initially landed on it, was reduced by 18ms per character. Furthermore, the gaze cursor also significantly reduced the character-to-character variance of T_{click} but not for T_{shift} . While a duration reduction of the click phase was expected, it is interesting to note that the duration of gaze shift also significantly reduced. Cursor stability, confidence, and fewer errors brought by the gaze-lock cursor may be the main reasons for the acceleration of the click phase. Participants performing with the gaze-lock cursor also benefited from better typing rhythm as demonstrated by the significant reduction of character-to-character variance. Once the typing rhythm was established, the gaze shift time could also be reduced due to the fluent typing process.

Compared to conventional gaze cursor, the gaze-lock cursor was designed with the addition of a spatial constraint and two temporal constraints, which contributed to the performance improvement. Once the gaze point stayed in target key for 100ms (T_{lock}), the displayed gaze cursor would be locked at the key center and waited for a manual click. This spatial constraint reduced the jitters of the displayed gaze cursor and improved the participant's confidence of manually clicking, thus speeding up the action. In contrast, the conventional gaze cursor exhibited more jittery motions even though the participants tried to fixate on the target. To make the problem even worse, people would involuntarily try to neutralize the position deviation caused by eye jitters and measurement noise of the eye tracking system, thus slowing down their manual click action. In addition, the gaze-lock cursor would keep the cursor locked until the gaze point had moved off the key for over 100ms (T_{unlock}). This unlock temporal constraint could reduce selection failures or errors that occurred when the user's gaze moved out of the desired target key at the time of manual click. This type of error occurs due to two possible reasons. Firstly, the gaze cursor may fall out of the target due to eye jitters and/or measurement noise of the eye tracking system (Kumar et al. 2008). Secondly, it was shown that people tended to move their gaze away before manual actions when using desktop applications with keyboard and mouse (Aoki, Hansen, and Itoh 2008; Liebling and Dumais 2014). By applying the temporal constraint for unlocking, the gaze-lock method provided a promising solution to mitigate this eye-hand coordination problem.

While the temporal constraint for unlocking was beneficial for gaze interactions, it brought in latency when the user intended to move the gaze away. In that case, the displayed cursor would not track the user's gaze immediately but jumped to the gaze position after a short interval (i.e. T_{unlock}). Even though participants might find this odd

at the beginning, they appeared to get used to it after a short time of training. In the formal test, none of our participants reported that they felt a delay of cursor control with the gaze-lock cursor. Instead, a common remark of the participants about the gaze-lock cursor, after experiencing both types of gaze cursor, was that “I feel that the cursor controlled by my gaze is more stable and accurate”.

Our study demonstrated that better performances were achieved with larger key width. This trend has been observed in previous studies (Kim et al. 2017; Zhang, Ren, and Zha 2008), and is in accordance with the prediction of Fitts’ law (Soukoreff and MacKenzie 2004). Moreover, we found that the larger key gap also led to significant improvement in performance, although a previous study reported a null effect of target gap during gaze interaction (Zhang, Ren, and Zha 2008). This inconsistency might result from the differences in the gaze interaction approach and the key gap employed. The previous research used the dwell approach to perform a key selection, while the present study used the eye-hand combination approach. The other possible reason is that Zhang and colleagues used much larger key gaps than the current study: their smallest key gap was 0.94° while we used 1.2° and 0.2° . Thus, it is also possible that the key gap might have had a ceiling effect in Zhang et al.’s study.

Another interesting trend observed in our data was that all performance measures improved with the increase of key-to-key distance (the sum of key width and key gap). In addition, the two conditions with the same key-to-key distance (1.2° & 1.2° vs. 2.2° & 0.2°) showed similar performance (Figure 3-7), which is not in accordance with the prediction of Fitts’ law. This implies that the key-to-key distance rather than the key width or the key gap alone is a better predictor of the eye-hand combination input performance when targets are tightly arranged.

We found significant interactions between gaze cursor mode and key width for input speed, error rate, and edit distance, between gaze cursor mode and key gap for error rate and edit distance. In general, these interactions demonstrate that the advantage of the gaze-lock cursor over the conventional gaze cursor is more evident in relatively challenging task conditions, i.e. typing with smaller key width/gap. This is practically important for human computer interactions, considering that small key widths and gaps are quite common in the general user interfaces (GUIs) of many kinds of software. For instance, the “Start” key in the Microsoft Word 2013 software (Chinese version) is no more than 1.2cm on a 23-inch display, and the average gap between two adjacent functional buttons on the menu bar is no more than 0.5cm. To accommodate these kinds

of small targets, previous research on the gaze-based target selection proposed several methods, including the use of a fish-eye lens for magnification (Ashmore, Duchowski, and Shoemaker 2005), displaying a magnified view of the region that the user is looking at when a key is pressed (Kumar, Paepcke, and Winograd 2007), and expanding the target area upon recognition of user's intent to interact with the target (Miniotas, Spakov, and Mackenzie 2004). The performance improvement of the gaze-lock cursor in dealing with small targets suggests that it is a promising method for various applications, especially those with crowded target items.

There are several limitations in the present study. First, we only compared the gaze-lock cursor with the conventional gaze cursor. Future studies should include a comparison with other methods aimed at stabilizing the gaze cursor, such as the ALCM method or the "force field" gaze data filtering method (Zhang, Ren, and Zha 2008), to obtain a better understanding of their relative advantages and disadvantages. Second, our experiments were performed only by young and middle-aged adults (age range 22-37 years). As previous research have suggested that age group is an important factor influencing human computer interactions (Jochems, Vetter, and Schlick 2013; Murata 2006), future research should collect data of older adults and children to make the conclusions more generalized. Last but not least, the key parameters for our gaze-lock method, the locking and unlocking time thresholds (T_{lock} and T_{unlock}), were both set to 100ms according to our piloting experiments. Future studies should adopt certain adaptive methods to dynamically set the time thresholds for different tasks and for different users, in order to optimize the task performance and user experience.

Acknowledgement

The study was supported by the National Key Research and Development Program of China (No. 2017YFB1300305), Foundation of National Key Laboratory of Human Factors Engineering (No. 6142222180305, SYFD180051802), and National Natural Science Foundation of China (No. 31622029, 31671168, 61533001). We thank Xinyong Zhang, Erez James Cohen and two anonymous reviewers for their constructive suggestions for the improvement of the manuscript.

References

- Aoki, Hirotaka, John Paulin Hansen, and Kenji Itoh. 2008. "Learning to interact with a computer by gaze." *Behaviour & Information Technology* 27 (4):339-44.
- Ashmore, Michael, Andrew T Duchowski, and Garth Shoemaker. 2005. Efficient eye pointing with a fisheye lens. Paper presented at the Proceedings of Graphics interface 2005.
- Bates, Richard, Stephen Vickers, and Howell O Istance. 2010. "Gaze interaction with virtual on-line communities: levelling the playing field for disabled users." *Universal Access in the Information Society* 9 (3):261-72.
- Blattgerste, Jonas, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. Paper presented at the Proceedings of the Workshop on Communication by Gaze Interaction.
- Bulling, Andreas, and Hans Gellersen. 2010. "Toward Mobile Eye-Based Human-Computer Interaction." *IEEE Pervasive Computing* 9 (4):8-12.
- Chin, Craig A, Armando Barreto, J Gualberto Cremades, and Malek Adjouadi. 2008. "Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities." *Journal of Rehabilitation Research and Development* 45 (1):161-74.
- Cho, Dongchan, Wahseng Yap, Heekyung Lee, Injae Lee, and Whoiyul Kim. 2012. "Long range eye gaze tracking system for a large screen." *IEEE Transactions on Consumer Electronics* 58 (4):1119-28.
- Choe, Mungyeong, Yeongcheol Choi, Jaehyun Park, and Hyun Kwang Kim. 2019. "Comparison of Gaze Cursor Input Methods for Virtual Reality Devices." *International Journal of Human-Computer Interaction* 35 (7):620-9.
- De Gaudenzi, Enrico, and Marco Porta. 2013. Towards effective eye pointing for gaze-enhanced human-computer interaction. Paper presented at the science and information conference.
- Duchowski, Andrew T. 2018. "Gaze-based interaction: A 30 year retrospective." *Computers & Graphics* 73:59-69.
- He, Jibo, Alex Chaparro, Bobby Nguyen, Rondell J Burge, Joseph Crandall, Barbara Chaparro, Rui Ni, and Shi Cao. 2014. "Texting while driving: Is speech-based text entry less risky than handheld text entry?" *Accident Analysis & Prevention* 72:287-95.
- Jacob, Robert J K. 1991. "The use of eye movements in human-computer interaction techniques: what you look at is what you get." *ACM Transactions on Information Systems* 9 (2):152-69.
- Jochems, Nicole, Sebastian Vetter, and Christopher Schlick. 2013. "A comparative study of information input devices for aging computer users." *Behaviour & Information Technology* 32 (9):902-19.
- Kim, Minjee, Daehwan Jin, Ilsun Rhiu, and Myung Hwan Yun. 2017. "The Effect of Stimulus Size and Position on the Task Performance of an Eye Mouse: Comparing Blink and Dwell Methods in a Click Task." *International Journal of Human-Computer Interaction* (3):1-18.
- Klamka, Konstantin, Andreas Siegel, Stefan Vogt, Fabian Gobel, Sophie Stellmach, and Raimund Dachselt. 2015. Look & Pedal: Hands-free Navigation in Zoomable Information Spaces through Gaze-supported Foot Input. Paper presented at the international conference on multimodal interfaces.

- Kumar, Manu, Tal Garfinkel, Dan Boneh, and Terry Winograd. 2007. Reducing shoulder-surfing by using gaze-based password entry. Paper presented at the symposium on usable privacy and security.
- Kumar, Manu, Jeff Klingner, Rohan Puranik, Terry Winograd, and Andreas Paepcke. 2008. Improving the accuracy of gaze input for interaction. Paper presented at the eye tracking research & application.
- Kumar, Manu, Andreas Paepcke, and Terry Winograd. 2007. EyePoint: practical pointing and selection using gaze and keyboard. Paper presented at the human factors in computing systems.
- Kurauchi, Andrew, Wenxin Feng, Carlos Morimoto, and Margrit Betke. 2015. HMAGIC: head movement and gaze input cascaded pointing. Paper presented at the Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments.
- Liebling, Daniel J, and Susan T Dumais. 2014. Gaze and mouse coordination in everyday work. Paper presented at the Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication.
- Majaranta, Päivi, and Andreas Bulling. 2014. "Eye Tracking and Eye-Based Human-Computer Interaction." In *Advances in Physiological Computing*, edited by Stephen H. Fairclough and Kiel Gilleade, 39-65. London: Springer London.
- Meister, Daniel. 2015. "Using swaps and deletes to make strings match." *Theoretical Computer Science* 562:606-20.
- Miniotas, Darius, Oleg Spakov, and I Scott Mackenzie. 2004. Eye gaze interaction with expanding targets. Paper presented at the human factors in computing systems.
- Miniotas, Darius, Oleg Spakov, Ivan Tugoy, and I Scott Mackenzie. 2006. Speech-augmented eye gaze interaction with small closely spaced targets. Paper presented at the eye tracking research & application.
- Miyoshi, Tetsuya, and Atsuo Murata. 2001. Input device using eye tracker in human-computer interaction. Paper presented at the robot and human interactive communication.
- Mott, Martez E., Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. 2017. Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times. Paper presented at the Chi Conference on Human Factors in Computing Systems.
- Murata, Atsuo. 2006. "Eye-Gaze Input Versus Mouse: Cursor Control as a Function of Age." *International Journal of Human-Computer Interaction* 21 (1):1-14.
- Murata, Atsuo, and Waldemar Karwowski. 2019. "Automatic Lock of Cursor Movement: Implications for an Efficient Eye-Gaze Input Method for Drag and Menu Selection." *IEEE Transactions on Human-Machine Systems* 49 (3):259-67.
- Pinheiro, Carlos G, Eduardo Lm Naves, Pierre Pino, E Losson, Adriano O Andrade, and Guy Bourhis. 2011. "Alternative communication systems for people with severe motor disabilities: a survey." *Biomedical Engineering Online* 10 (1):31-.
- Soukoreff, R William, and I Scott MacKenzie. 2004. "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI." *International Journal of Human-Computer Studies* 61 (6):751-89.
- Stawicki, Piotr, Felix Gembler, Aya Rezeika, and Ivan Volosyak. 2017. "A Novel Hybrid Mental Spelling Application Based on Eye Tracking and SSVEP-Based BCI." *Brain Sciences* 7 (4):35.

- Stellmach, Sophie, and Raimund Dachzelt. 2012. Look & touch: gaze-supported target acquisition. Paper presented at the human factors in computing systems.
- Tian, Yu, Yifei Huang, Jibo He, and Kunlin Wei. 2018. "What affects gait performance during walking while texting? A comparison of motor, visual and cognitive factors." *Ergonomics*:1-12.
- Tuisku, Outi, Veikko Surakka, Ville Rantanen, Toni Vanhala, and Jukka Lekkala. 2013. "Text entry by gazing and smiling." *Advances in Human-computer Interaction* 2013:1-13.
- Tuisku, Outi, Veikko Surakka, Toni Vanhala, Ville Rantanen, and Jukka Lekkala. 2012. "Wireless Face Interface: Using voluntary gaze direction and facial muscle activations for human-computer interaction." *Interacting with Computers* 24 (1):1-9.
- Tula, Antonio Diaz, Andrew T N Kurauchi, and Carlos Hitoshi Morimoto. 2013. Facilitating gaze interaction using the gap and overlap effects. Paper presented at the human factors in computing systems.
- Wang, Jingtao, Shumin Zhai, and Hui Su. 2001. Chinese input with keyboard and eye-tracking: an anatomical study. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Wobbrock, Jacob O, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. Paper presented at the human factors in computing systems.
- Yu, Chun, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture?: Exploring Head-Based Text Entry Techniques for HMDs. Paper presented at the human factors in computing systems.
- Yu, Mingxin, Yingzi Lin, David K Schmidt, Xiangzhou Wang, and Yu Wang. 2014. "Human-Robot Interaction Based on Gaze Gestures for the Drone Teleoperation." *Journal of Eye Movement Research* 7 (4).
- Zhai, Shumin, Carlos Morimoto, and Steven Ihde. 1999. Manual and gaze input cascaded (MAGIC) pointing. Paper presented at the Proc. Human Factors in Computing Systems: CHI 99 Conference.
- Zhang, Xinyong, Wenxin Feng, and Hongbin Zha. 2011. Effects of different visual feedback forms on eye cursor's stabilities. Paper presented at the international conference on internationalization design and global development.
- Zhang, Xinyong, Xiangshi Ren, and Hongbin Zha. 2008. Improving eye cursor's stability for eye pointing tasks. Paper presented at the human factors in computing systems.