Eyes and Keys: An Evaluation of Click Alternatives Combining Gaze and Keyboard

Ken Neth Yeoh, Christof Lutteroth, Gerald Weber

Department of Computer Science, University of Auckland,
38 Princes Street, Auckland 1010, New Zealand
kyeo475@aucklanduni.ac.nz {christof,gerald}@cs.auckland.ac.nz

Abstract. With eye gaze tracking technology entering the consumer market, there is an increased interest in using it as an input device, similar to the mouse. This holds promise for situations where a typical desk space is not available. While gaze seems natural for pointing, it is inherently inaccurate, which makes the design of fast and accurate methods for clicking targets ("click alternatives") difficult. We investigate click alternatives that combine gaze with a standard keyboard ("gaze & key click alternatives") to achieve an experience where the user's hands can remain on the keyboard all the time. We propose three novel click alternatives ("Letter Assignment", "Offset Menu" and "Ray Selection") and present an experiment that compares them with a naive gaze pointing approach ("Gaze & Click") and the mouse. The experiment uses a randomized, realistic click task in a web browser to collect data about click times and click accuracy, as well as asking users for their preference. Our results indicate that eye gaze tracking is currently too inaccurate for the Gaze & Click approach to work reliably. While Letter Assignment and Offset Menu were usable and a large improvement, they were still significantly slower and less accurate than the mouse.

Keywords: Eye gaze tracking, click alternative, keyboard.

1 Introduction

Pointing is a natural activity of the human eye: when we look at an object, this is a good indicator that the object is currently occupying our attention; this rule of thumb has become a core principle in research based on gaze tracking. As a result, gaze tracking is a promising technology for natural user interfaces. It does not require learning any new techniques. With eye gaze tracking devices entering the consumer market at prices similar to gaming mice, there is a growing interest among users as well as in the HCI community.

Gaze trackers hold promise for a variety of situations. They have become an established assistive technology for improving accessibility. While gaze tracking has yet to become a widely used technology, it holds promise for work away from the desk, especially with laptops. Laptops can provide a keyboard similar to desktop devices, but they have no adequate surface for a normal mouse, e.g. when used on the lap. Furthermore, the use of a mouse in typical productive work requires users to switch a hand between keyboard and mouse frequently, which incurs a time penalty. The use of gaze for pointing would allow users to keep their hands on the keyboard more persistently. In any case, users must identify a target visually before moving the mouse to click it, which means that eye gaze tracking could lead to more direct and fluent interaction. Gaze tracking technology may also mitigate some of the problems of repetitive strain injury (RSI) related to mouse overuse.

While pointing with the eyes may seem easy from a user's perspective, eye gaze tracking technology inherently suffers from inaccuracy. First, there are technical challenges of calibration, resolution, tracking volume, changing lighting conditions, variances in the anatomy of the face and eyes, and optical properties of visual aids such as glasses and contact lenses. All these factors make it difficult to achieve a good tracking accuracy for all users. Second, there are physiological limitations: involuntary eye movements such as jitter and drifts, blinks, and the size of our fovea which also provides clear vision of objects that are not exactly on the point of gaze. These challenges and limitations make it difficult to design methods for pointing and clicking UI elements based on gaze (“click alternatives”) that are fast and accurate.

In this research we investigate click alternatives that make use of eye gaze tracking for pointing and a standard keyboard for clicking (“gaze & key click alternatives”). This fills a gap in the current literature, which mostly focuses on purely gaze-based or gesture-based click alternatives using techniques such as zooming, dwell time thresholds and confirm buttons. We propose three novel gaze & key click alternatives (“Letter Assignment”, “Offset Menu” and “Ray Selection”), specify them formally using state machines and discuss their underlying design decisions. In an experimental study we compare the click alternatives with naive gaze-pointing and keyboard clicking (“Gaze and Click”) and the mouse. We address the following research questions:

RQ1. How can fast and accurate gaze & key click alternatives be designed?
RQ2. How do gaze & key click alternatives compare with each other and the mouse?
RQ3. Are gaze & key click alternatives mature enough for everyday general use?

Our results shed light on the problems and opportunities of gaze & key click alternatives, providing guidance for interaction designers of gaze-based user interfaces. Furthermore, we share information and resources that will help other researchers perform realistic click alternative evaluations. In summary, this paper makes the following contributions:

- Three novel gaze & key click alternatives with formal state machine specifications.
- An experimental comparison of gaze & key click alternatives, with insights into the interaction design and general usability of such technologies.
- An experimental procedure for the evaluation of click alternatives and an open-source implementation of the aforementioned gaze & key click alternatives.

Section 2 summarizes related work about gaze-based click alternatives. Section 3 describes the design of the proposed gaze & key click alternatives. Section 4 describes the experimental methodology used. Section 5 gives an overview of the results, and Section 6 discusses them. Section 7 summarizes conclusions and points out future research directions.

2 Related Work

Eye gaze tracking as a pointing device currently lacks the accuracy required to be used as a simple point & click device [1-5], for a number of different reasons. Firstly, the fovea of the eye, which is responsible for sharp central vision, covers about one degree of visual angle [4,6]. This relatively large angle means that it may be difficult for the eye gaze tracker to accurately pinpoint what the user is looking at on the screen, especially if the target is small such as an icon or text. At a distance of 65cm from the screen, the eye can view an area of about 1.1cm diameter clearly. Furthermore, our gaze subconsciously drifts or jumps to other points of interest. As a result, it takes a conscious effort from the user to hold the gaze in an area for a length of time [5]. These eye gaze tracking inaccuracies cannot be solved by simply upgrading the hardware; therefore different software solutions have been built to increase the accuracy in pinpointing the user's gaze.

Zhang et al. [7] proposed techniques to improve the stability of an eye gaze cursor, using force fields, speed reduction and warping to a target’s center. Force fields act as a kind of magnet for the cursor: the algorithm attempts to deduce the user's intent and tries to prevent the cursor from veering off target. Cursor speed reduction was found to increase speed and accuracy when using the eye gaze pointer for medium-size targets. Such techniques are useful, but do not improve the accuracy of eye gaze cursors sufficiently for general use.

The most obvious and natural purely gaze-based click alternative is “dwell”, which clicks a target after the gaze dwells on it for a certain time. For simple object selection tasks, dwell can be significantly faster than the mouse [8]. However, while it has been successfully used for specialized UIs such as carefully-designed menus [9,10], dwell alone is generally insufficient as a general click alternative because it is not accurate enough for small targets. Hardware buttons for clicking seem slightly faster than simple dwell with a typical 0.4 seconds dwell threshold, but less accurate as people tend to click before the gaze has fully settled on the target [11]. The accuracy can be improved by taking into account system lag and delaying triggers accordingly [12].

One approach to address the lack of accuracy is to enlarge or zoom in on the general area of the user's gaze. EyePoint [13] magnifies the area around the gaze when a hotkey is pressed, and performs a click at the point of gaze in the magnified view when the hotkey is released. The reported click times are fast (below 2 seconds), but there are problems with accuracy (error rate exceeding 10%). There are similar techniques relying only on gaze: Zoom Navigator [3] continuously magnifies the area the user is looking at, until it is clear what the target is and the target is automatically clicked. If correcting movements are made, Zoom Navigator zooms out for a short period before continuing to zoom in again.

Most zooming techniques overlay the area under the cursor with a magnification, so context is lost. There are techniques to mitigate the loss of contextual information, such as fish-eye lenses and offset magnifying glasses. Ashmore et al. [14] investigated a dwell-activated fish-eye lens with a continuous fish-eye zoom, which preserves but distorts the context around the target. FingerGlass [15] employs an offset magnifying glass which never covers the zoomed-in area for touch-based interaction. However, when applying offset techniques in gaze-based interaction, one must consider that the offset content will immediately attract the user's gaze.

Bates et al. [4] investigated gaze clicking with zooming and found a clear relationship between the target size and the level of magnification used by a user when targeting a small area. Participants would zoom in until the target is just larger than the pre-test measured pointing accuracy of the eye gaze tracker. It was also found that the participants had difficulty maintaining focus on a target during the selection process. The time spent correcting the cursor position on targets was the largest portion of non-productive time spent carrying out the tasks. This emphasizes the need for additional techniques to address the inaccuracy of gaze cursors.

ceCursor [2] uses transparent directional buttons located around the area the user is looking at. The buttons, which are activated by dwelling on them, can be used to move a cursor. This technique is accurate (even for small targets) but slow, taking on average 11.95 seconds. Using a keyboard, it would be straightforward to use the directional keys in a similar manner. But while accurate, this would be slow compared to the speed of gaze.

The gaze-based WeyeB browser [1] uses a combination of dwell and eye gestures for link navigation. Once the user is looking at the desired target, they must flick their eyes upwards and then back downwards to click a link. If multiple links are under the general area of the cursor, a large secondary drop down menu with the different link options is displayed – an alternative to zooming. The combination of dwell and eye gestures solved the "Midas touch" problem, i.e. inadvertent clicking that can occur when using dwell alone. Gaze & key click alternatives generally do not suffer from the Midas touch, as a key can be used to clearly signal a click.

Another method of improving the accuracy of pointing with eye gaze is to use facial movements to refine the cursor position [16]. Four electrodes are placed on the user in order to capture electromyogram (EMG) signals from muscles in the face. The user first looks at the approximate target location, then uses facial movements to incrementally move the cursor, and finally performs click actions using other facial movements. While this increased accuracy to near mouse levels, it was still about four times slower (more than 4 seconds per click).

Some approaches combine gaze tracking with a physical pointing device. MAGIC [17] moves the pointer quickly to the gaze position to speed up pointing, using the mouse for finer movements and clicking. The Rake Cursor [18] shows a grid of multiple mouse pointers simultaneously, moving the whole grid with the mouse and selecting the active pointer in the grid by gaze. It successfully reduces mouse movements as the pointer closest to a target can be used. The Gaze-enhanced User Interface Design (GUIDe) [19] combines gaze with keyboard and mouse to improve various common tasks.

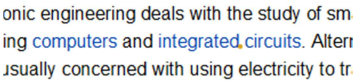
3 Click Alternative Design

Four gaze & key click alternatives were designed and implemented as follows. The Ray Selection alternative is included here, but was not used in the experiment for reasons outlined later in this section. All click alternatives are designed with a web browser as the basis, so the targets in the following examples are hyperlinks. The click alternatives can also be applied to other types of targets. All click alternatives were implemented in Java using the Webkit[1] web browser engine as a basis. They are freely available as open-source software[2].

3.1 Gaze & Click

This is the simplest of the four click alternatives: gaze is used for pointing and a hotkey (we chose the ‘F’ key) is used for clicking. If the hotkey is pressed while the recorded gaze position is directly on a link, then the link is clicked. Otherwise, no link is clicked. Figure 1 shows a screenshot of the implementation. Figure 2 illustrates the overall interaction in a state machine diagram, using the notation from [20]. Its modeless nature is clearly expressed in the presence of a single state.

As indicated by the related work, it is very hard to use this click alternative, and this was confirmed in our pilot study. In particular, it was simply too difficult for the users to know if a link is currently underneath the recorded gaze position. It was necessary to add a visual gaze cursor, an orange dot representing the user’s current gaze position. While it may be possible to hide the gaze cursor for larger targets, most textual hyperlinks are simply too small given the typical inaccuracy of gaze tracking. With a visible gaze cursor users are at least aware of the gaze tracking error and can compensate for it by adjusting their gaze.



onic engineering deals with the study of sm
ing computers and integrated circuits. Altern
usually concerned with using electricity to tr

Fig. 1. Screenshot of the Gaze & Click implementation (note the gaze cursor after “integrated”)

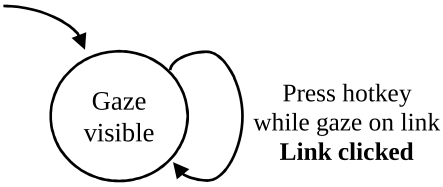


Fig. 2. State machine of the Gaze & Click alternative

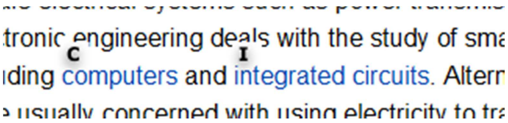
[1] http://www.webkit.org/

[2] http://github.com/aucklandhci/gazebrowser

3.2 Letter Assignment

The main idea of the Letter Assignment alternative is to assign letters to links close to the gaze and allow the user to choose between them by pressing the corresponding key. This is similar to hotkeys, with the difference that assigned letters are shown near the gaze and both gaze and keys are used for disambiguating targets. The use of gaze ensures that all visible links are clickable with a single keystroke even on crowded pages. Figure 3 shows a screenshot of the implementation. Figure 4 shows the state machine diagram. After gazing at an area with at least one hyperlink, a unique capitalized letter is drawn slightly above all hyperlinks within a radius around the gaze position. The letter shown for each hyperlink is chosen as the first available character in the label of the link, making it easier for users to anticipate the letter for a link. If there is no unique character in the hyperlink's label, the next available letter in the alphabet is used instead. To click the hyperlink, the user presses the corresponding key.

A white rectangle is drawn behind the overlaid letter to allow easier reading of the letter; since the hyperlinks are often quite close to other text, the overlaid letters could otherwise be hard to make out. A drop shadow is drawn behind the overlaid letters, on top of the white rectangle, to give the illusion of layering; the overlaid letters are on top and the web browser is the background. The user will naturally want to interact with the top-most layer. The color of the letters is kept black (the same as most of the text on the page) to make them less distracting, so they can be ignored more easily if clicking hyperlinks is not the user's intention.



tronic engineering deals with the study of sma
iding computers and integrated circuits. Altern
e usually concerned with using electricity to tra

Fig. 3. Screenshot of the Letter Assignment implementation

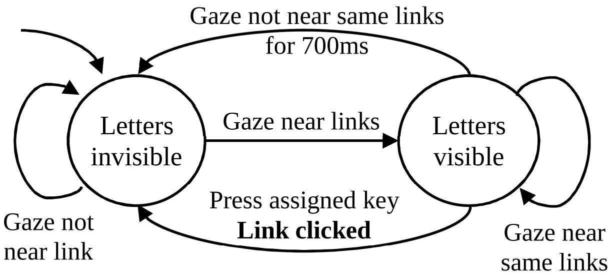
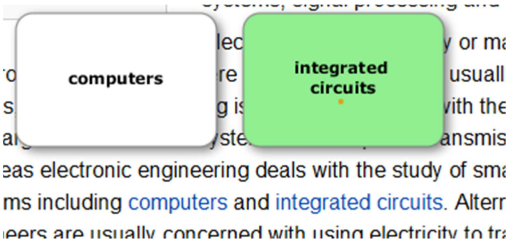


Fig. 4. State machine of the Letter Assignment click alternative



eas electronic engineering deals with the study of sma
ms including computers and integrated circuits. Altern

Fig. 5. Screenshot of the Offset Menu implementation

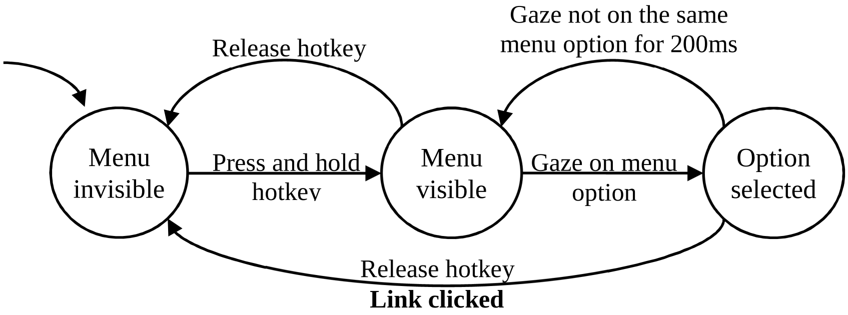


Fig. 6. State machine of the Offset Menu click alternative

3.3 Offset Menu

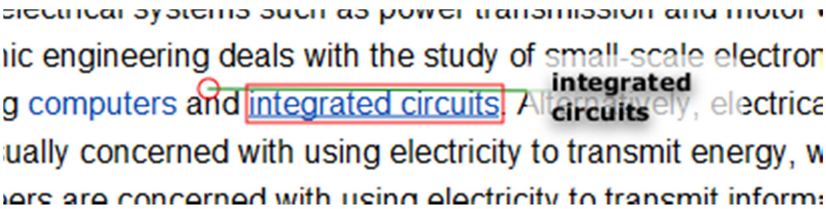
The main idea of the Offset Menu is similar to that of confirm buttons [20], i.e. buttons that are used to confirm a click action after pointing with gaze. Instead of dwell-activated confirm buttons in a page margin, a menu with buttons is activated with a hotkey (we chose the ‘F’ key again). Because the menu cannot appear accidentally, as with the dwell-activated confirm buttons, the menu can be placed at an offset to the gaze position without disruption. A screenshot of the implementation is shown in Fig. 5 and the state machine diagram in Fig. 6. When the user presses and holds the hotkey, a menu is displayed at an offset above the current gaze position. If the offset menu would be cut off by any screen edge, it is shifted to display correctly. The menu contains an entry for each hyperlink within a radius of the reported gaze, i.e. the same set of hyperlinks as in Letter Assignment. If there are no hyperlinks in the vicinity, no menu is drawn. While still holding down the hotkey, the user can shift their gaze to one of the menu options, and the menu option will turn green to show it is currently selected. The user can then release the hotkey, and whichever menu option is currently selected will be clicked. If no menu option is selected when the hotkey is released, no hyperlink will be clicked. There is a de-selection threshold of 200ms, to prevent jitter in the gaze coordinates from unintentionally de-selecting an option.

A drop shadow is again drawn behind the menu options to create the illusion of layering; the menu options are on top and the web page is in the background. The selected option is green because the color green affords “going forward”, much like a traffic light. The size of the menu options is large enough to allow for some inaccuracy and imprecision in gaze tracking. The text is centered, drawing the user's attention to the center of the menu option to make it easier to select.

3.4 Ray Selection

The main idea of this alternative is to disambiguate links by selecting the direction of the intended target, similar to a radial menu. Figure 7 shows a screenshot of the implementation. Figure 8 shows the state machine diagram. When the user presses and holds the hotkey (we chose the 'F' key again), a ray is drawn from the recorded gaze point at the time the hotkey was pressed (the start point) towards the current gaze point. This ray is redrawn as the current gaze point changes. The ray may intersect with hyperlinks. The intersected hyperlink closest to the current gaze point is the selected hyperlink; it is highlighted with a red border and its name is drawn at the end of the ray, at the user's current gaze position, to let the user know which hyperlink is currently selected. When the user releases the hotkey, the selected hyperlink is clicked. If no hyperlink intersects the ray, no link is clicked.

The selected hyperlink's name is drawn on a semi-opaque white background, so it is possible to read the name even against background text. A drop shadow creates the illusion of layers, similar to Letter Assignment and Offset Menu. The selected hyperlink is highlighted with a red border to make it clear to the user which hyperlink is selected, even if the user is not looking at it directly.



electrical systems such as power transmission and motor

nic engineering deals with the study of small-scale electron

g computers and integrated circuits. Alternatively, electrica

ually concerned with using electricity to transmit energy, w

ers are concerned with using electricity to transmit inform

integrated circuits

Fig. 7. Screenshot of the Ray Selection implementation

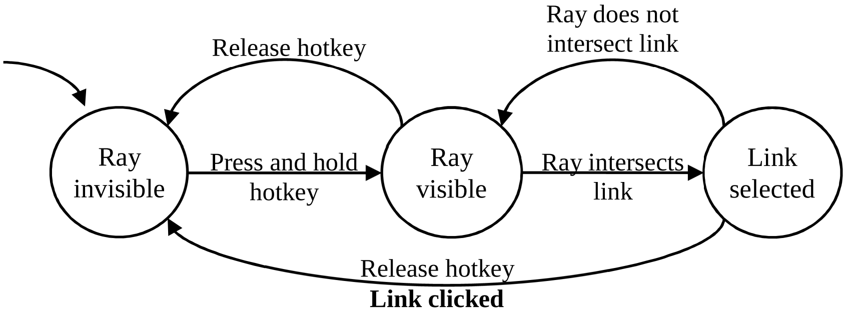


Fig. 8. State machine of the Ray Selection click alternative

During our pilot study, it became clear that this alternative exhibited several clear disadvantages, to a degree that it was clearly not worth to include it into the main study. We present this alternative in the interest of reporting also negative results, so that others can learn from our experience. Users often found it very difficult to click the desired hyperlink as they were unsure where to look to select a target. In particular, they found it difficult to change the currently selected hyperlink if it was not the desired one. The start point would typically already be close to the target (as in Fig. 7), forcing users to look beyond the target to make the ray point into the right direction despite gaze tracking inaccuracy. Our experience with this alternative illustrates the problems of separating gaze from intention, i.e. of making users look at anything that is not clearly a target.

4 Methodology

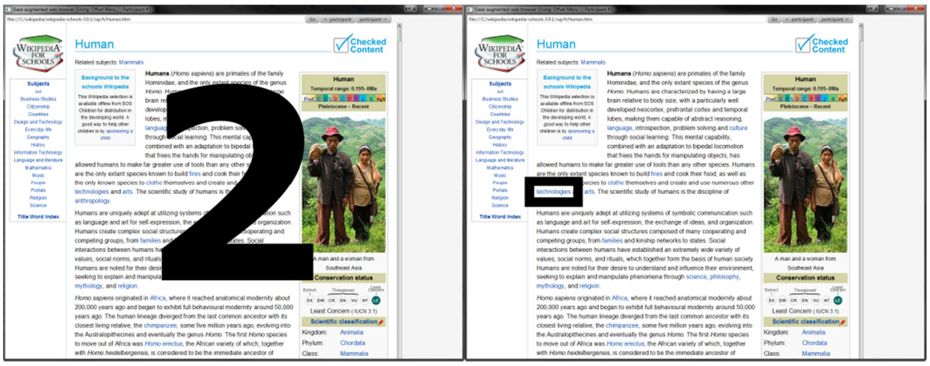
The usability study was conducted using a within-subjects design to reduce error variance stemming from individual performance differences. The independent variable is the click alternative used to complete the given tasks. The dependent variables measured are “time taken to click link” (click time) and “number of incorrect clicks” (inaccuracy). Ease of use was measured using the System Usability Scale (SUS) [21].

A 30 inch 144 Hz LCD monitor with a resolution of 1920x1080 pixels, a standard QWERTY keyboard and standard mouse with the default Win7 configuration were used. A Tobii X2-30W eye gaze tracker with a refresh rate of 30Hz was mounted on a tripod below the monitor in a non-intrusive space. A fully adjustable chair with headrest and armrests allowed participants with various heights to be well within the tracking volume of the gaze tracker, helped keep their head still and be overall comfortable for the duration of the experiment. The room was lit by fluorescent lights, and the blinds were closed to block sunlight from interfering with the eye gaze tracker.

After filling out a pre-experiment demographics questionnaire, the participants were comfortably seated and the chair adjusted to best fit the eye gaze tracker’s usable parameters. Before each click alternative was started, the eye gaze tracker was calibrated using Tobii's EyeX software, which took 20-30 seconds. Additional calibration was provided if the participant moved around too much or found the calibration to be too inaccurate. Calibration was then measured using a custom program which logged how close the gaze tracker coordinates were to each of nine on-screen calibration points.

A generic clicking task was used to measure click time and accuracy, in a series of 40 hyperlinks pseudo-randomly chosen from Wikipedia. We chose Wikipedia because it is one of the most visited websites and all the participants had used it in the past. An offline Wikipedia[3] was used to ensure all the pages were static and consistent. For each click alternative, participants were allowed as much “free-play” time as they wanted, so they could learn how the click alternative worked properly and get used to navigating Wikipedia pages.

[3] http://schools-wikipedia.org/



Human

Related subjects: Mammals

Background to the schools Wikipedia

This Wikipedia selection is available offline from SOS Children for distribution in the developing world. A good way to help other children is by sponsoring a child

Humans (*Homo sapiens*) are primates of the family Hominidae, and the only extant species of the genus *Homo*. Humans are characterized by having a large brain relative to body size, with a particularly well developed neocortex, prefrontal cortex and temporal lobes, making them capable of abstract reasoning, language, introspection, problem solving and culture through social learning. This mental capability, combined with an adaptation to bipedal locomotion that frees the hands for manipulating objects, has allowed humans to make far greater use of tools than any other species. Humans are the only extant species known to build fires and cook their food, as well as the only known species to clothe themselves and create and use numerous other technologies and arts. The scientific study of humans is the discipline of anthropology.

Humans are uniquely adept at utilizing systems of symbolic communication such as language and art for self-expression, the exchange of ideas, and organization. Humans create complex social structures composed of many cooperating and competing groups, from families and kinship networks to states. Social interactions between humans have established an extremely wide variety of values, social norms, and rituals, which together form the basis of human society. Humans are noted for their desire to understand and influence their environment, seeking to explain and manipulate phenomena through science, philosophy, mythology, and religion.

Homo sapiens originated in Africa, where it reached anatomical modernity about 200,000 years ago and began to exhibit full behavioural modernity around 50,000 years ago. The human lineage diverged from the last common ancestor with its closest living relative, the chimpanzee, some five million years ago, evolving into the Australopithecines and eventually the genus *Homo*. The first *Homo* species to move out of Africa was *Homo erectus*, the African variety of which, together with *Homo heidelbergensis*, is considered to be the immediate ancestor of

Human

Temporal range: 0.195–0Ma

Pleistocene – Recent

A man and a woman from Southeast Asia

Conservation status

Least Concern (IUCN 3.1)

Scientific classification

Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia

Fig. 9. Screenshots of the task: a brief countdown (left) followed by highlighting of the click target (right)

Participants were told to click the target hyperlinks as fast and accurately as they could. Before each click, a brief countdown was shown, and then the target link was highlighted with a thick black rectangle (Fig. 9) until a click was performed. The series of target hyperlinks was the same for each participant, as the same starting seed was used. When a participant accidentally clicked the wrong target, the browser would still navigate to the correct target for consistency. For each click alternative the first 10 clicks were training trials; only the following 30 clicks were used for later analysis. The order in which participants used the click alternatives was permuted to mitigate order bias and training effects.

All clicks were logged in a CSV file together with fine-grained events, such as the time the target was found and the time a button was pressed. After each click alternative was tested, a post-task questionnaire was filled out by the participant, which contained the ten SUS questions answered on a five-point Likert-scale. After completing the tasks using all four click alternatives, a post-experiment questionnaire was filled in, which asked the participants to rank the click alternatives from one to four, with one the best. An optional comment section allowed participants to explain their rankings and express their thoughts on each of the click alternatives.

5 Results

The experimental data set is available on the web[4]. 20 participants successfully took part in the main experiment (16 men and 4 women). There was one other participant for whom we were unable to calibrate the eye gaze tracker. The unsuccessful calibration was most likely due to a very high difference in his glasses' strength between the left and right eye. The participants were all aged between 19 and 45, with a variety of ethnicities and a range of disciplines including Computer Science, Engineering and Psychology. Participants indicated reading between 2 and 12 hours a day (median 5.5), and using the computer between 3 and 12 hours a day (median 8.0). Nine partic-

[4] http://github.com/aucklandhci/gazebrowser/tree/master/datasets/

ipants wore either glasses or contact lenses of varying strengths. In total, 2400 clicks were measured, recorded and analyzed (600 for every condition). Table 1 summarizes the results.

5.1 Performance

A one-way within-subjects ANOVA was conducted to test the effect of the click alternative on click time, showing a very significant effect (F(3, 1797)=46.97, p <0.0001). Paired samples t-tests with Holm correction were used to make post hoc comparisons between the conditions. There were significant differences between the click times of all conditions (p<0.001) except for Letter Assignment and Offset Menu (p=0.13). It is clear that Gaze & Click is the slowest and the mouse is the fastest click alternative. This is illustrated in more detail in the click time distributions in Figs. 10-13. The red line shows the cumulative percentage over all measured clicks.

Another one-way within-subjects ANOVA was conducted to test the effect of the click alternative on the number of correct clicks, showing a very significant effect (F(3, 72)=9.97, p <0.0001). Paired samples t-tests with Holm correction were used to make post hoc comparisons between the conditions. Similar to click time, there were significant differences between the numbers of correct clicks of all conditions (p<0.01) except for Letter Assignment and Offset Menu (p=0.10). It is clear that Gaze & Click is the least accurate and the mouse is the most accurate click alternative.

Participants clearly had problems clicking the correct links with Gaze & Click, and also some problems with Letter Assignment and Offset Menu. The click time distributions indicate that the gaze & key alternatives suffer from a fairly large number of outliers compared to the mouse. However, Letter Assignment and Offset Menu are a large improvement over Gaze & Click. Some of the typical comments on performance from the participants are as follows:

- “Letter assignment was quick and easy to use. Mouse beats the other two because it was far more accurate”
- “Mouse is the one I’m used to. Offset Menu was quick and accurate. Letter Assignment required whole keyboard. Gaze and click was super inaccurate”
- “Gaze & click with very accurate eye tracker would outperform the other two”
- “Offset: had to literally search the alternatives → cumbersome. Gaze and click: sometimes hard to hit target. Letter assignment: good but change of focus between keyboard and screen not ideal. Mouse: slow movement speed”

Table 1. Summary of results (click time in seconds)

	Gaze & Click	*Letter Assignment*	*Offset Menu*	*Mouse*
Click time mean	4.28	2.71	3.03	1.26
Click time std. dev.	7.17	3.94	3.24	0.47
Click time median	2.07	1.79	2.00	1.16
Click time median 95% CI	[1.84, 2.31]	[1.70, 1.88]	[1.92, 2.09]	[1.14, 1.20]
Incorrect clicks	35.33%	10.17%	6.5%	0.5%
SUS score mean	53.7	71.6	77.1	91.5
SUS score std. dev.	14.4	14.9	13.5	9.2
Rank mean	3.85	2.48	2.38	1.3

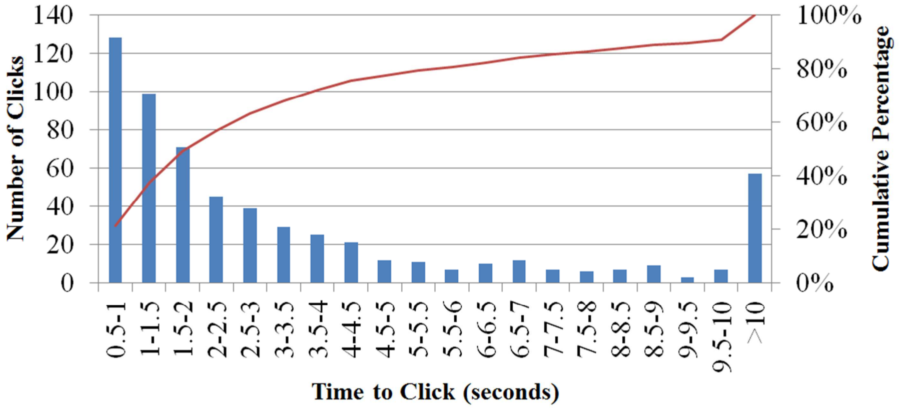


Fig. 10. Click time distribution for the Gaze & Click alternative

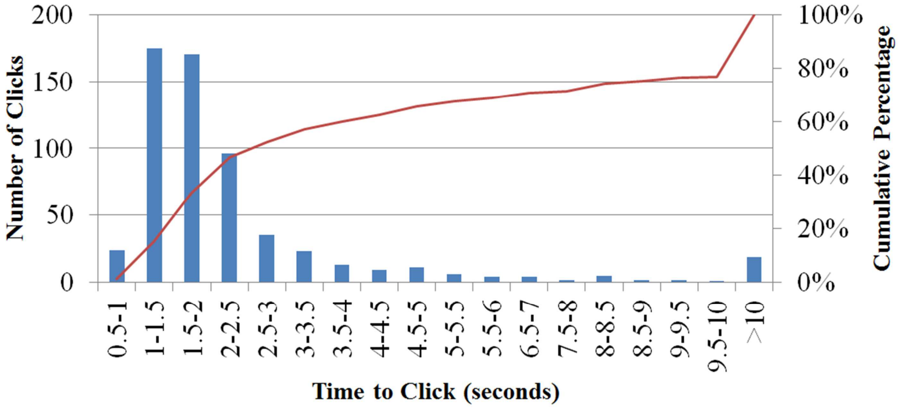


Fig. 11. Click time distribution for the Letter Assignment alternative

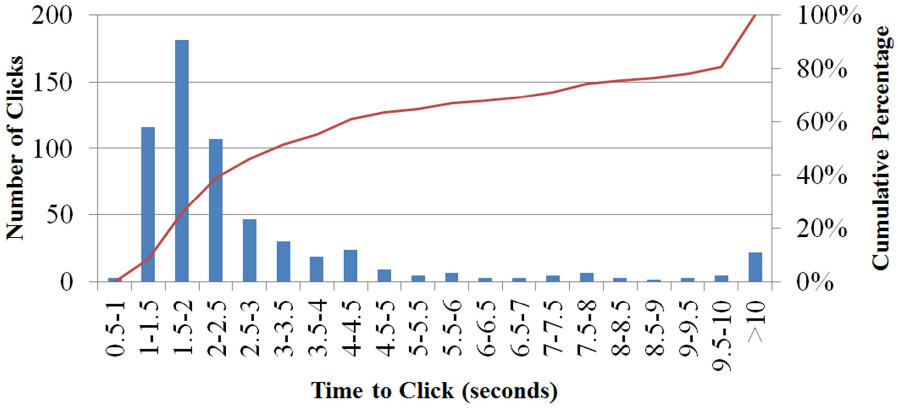


Fig. 12. Click time distribution for the Offset Menu alternative

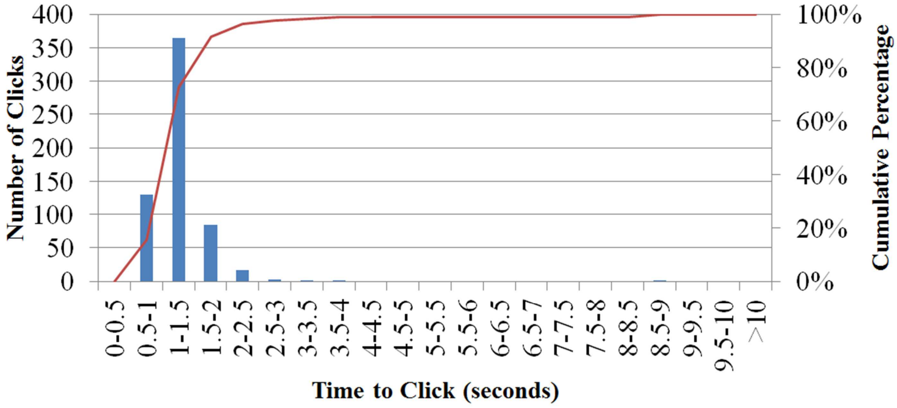


Fig. 13. Click time distribution for the mouse (note the beginning of the time-to-click axis at 0s as opposed to 0.5s in the previous figures)

5.2 Usability

Figure 14 shows the mean SUS scores of the click alternatives. A one-way within-subjects ANOVA was conducted to test the effect of the click alternative on SUS score, showing a very significant effect (F(3, 57)=39.13, p <0.0001). Paired samples t-tests with Holm correction were used to make post hoc comparisons between the conditions. Similar to the performance results, there were significant differences between the SUS scores of all conditions (p<0.001) except for Letter Assignment and Offset Menu (p=0.20). It is clear that Gaze & Click has the lowest and the mouse the highest SUS score.

Figure 15 shows a histogram of the click alternative rankings. A Friedman rank sum test was conducted to test the effect of the click alternative on rank, showing a very significant effect (χ2= 39.74, df=3, p<0.0001). Paired Wilcoxon signed-rank tests with Holm correction were used to make post hoc comparisons between the conditions. Consistent with the SUS scores, there were significant differences between the ranks of all conditions ($p<0.01$) except for Letter Assignment and Offset Menu ($p=0.82$). It is clear that Gaze & Click is ranked the worst and the mouse is ranked the best. Some typical comments on usability from participants are as follows:

- "Letter assignment was a little troublesome actually having to wait for and read the assigned letter, which is annoying when it's not the first letter"
- "Offset menu is predictable and lots of visual feedback"
- "Offset menu was easy to use but still seemed like a strain on the eyes with continuous use"

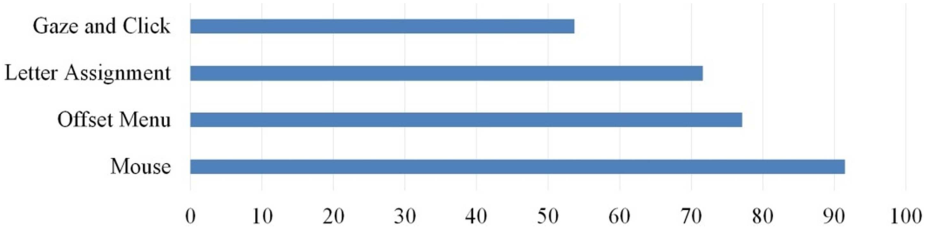


Fig. 14. Mean SUS scores of the click alternatives

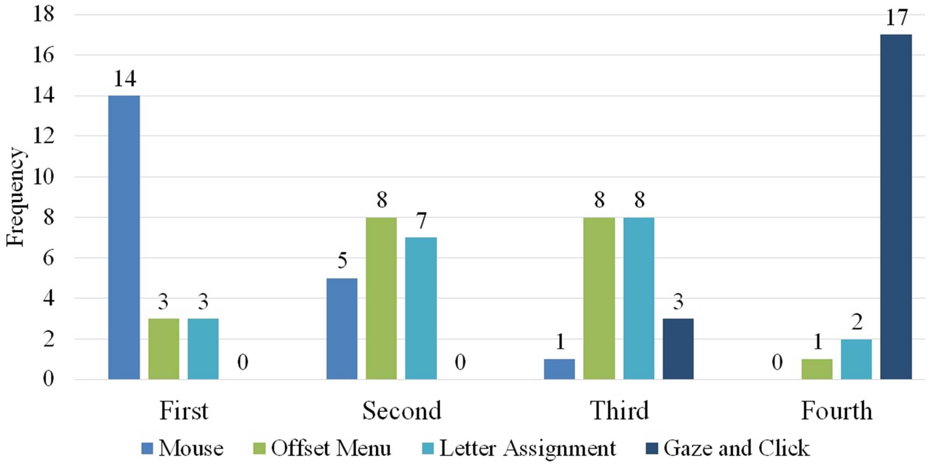


Fig. 15. Histogram of click alternative rankings

6 Discussion

As expected, the mouse was the fastest and most accurate click alternative, with the best SUS score. This may be in part due to every participant being heavily used to it and its easy learning curve. Gaze & Click has no support for helping users click links accurately; therefore we expected it to be the least accurate click alternative, but not necessarily the slowest. Interestingly, participants often went for one of two different approaches. One was to click links quickly regardless of whether they were sure their gaze position was on top of the right link, and the other was to spend a long time getting their gaze position to be stable on top of the right link before clicking the hotkey. The first approach was very quick, often being faster than the mouse, but the second approach sometimes took upwards of 10 seconds. Participants were told to "click links as fast and as accurately as possible", so participants had to decide whether to be fast or to be accurate, as it was clearly not possible to do both.

From observation, participants had trouble deactivating the Offset Menu if none of the options given were correct. There were two causes for this. First, the 200 ms deselection threshold was not explicitly explained to users beforehand. Second, some users looked too far off-screen, breaking the line of sight with the gaze tracker. This often caused the gaze position to freeze on a menu option. In both cases, the option would still be selected when the user released the key.

From observation, Letter Assignment proved to be difficult for participants because the assigned letter for a link was not always the one they were expecting. For example, two links "Citizenship" and "Countries" were often next to each other and both assigned a letter. "Citizenship" would be assigned the 'C' key and "Countries" would be assigned 'O'. Participants would often click 'C' if they wanted to go to "Countries". Most of the time the assigned letter was the one the participants were expecting, so it would trip them up when it was not.

For both Letter Assignment and Offset Menu, the speed was quite close to the mouse if the target link was among the first links selected (either by displaying the letters or showing the menu). However, if the target link was not immediately selected for whatever reason, then this would at least double the click time: the user would need to de-select and then re-select the options. This was not an issue for Gaze & Click or the mouse, as both of them are modeless and do not need a selection/deselection process.

How do the presented gaze & key click alternatives compare to other gaze-based click alternatives? It is difficult to compare the results of studies with different methodologies. However, some studies use hyperlink clicking tasks similar to the one presented here, so at least a discussion is possible. The purely gaze-based Multiple Confirm click alternative [20] seems slower than Letter Assignment and Offset Menu, which is not surprising considering that no hardware buttons are used. Interestingly, Multiple Confirm also seems more accurate, probably because it is harder to click incorrect (and correct) links. EyePoint [13], which is another gaze & key click alternative, seems faster than Letter Assignment and Offset Menu, but less accurate than Offset Menu. This could be because Offset Menu – in contrast to EyePoint – gives clear feedback about the target that will be clicked.

7 Conclusion

Eye trackers may well be one of the next types of computer peripherals going mainstream. However, it is still a challenge to create added value off these devices in everyday computing. Using them as a pointing device seems natural, and combining them with a keyboard to create a point-and-click interface may have advantages in situations where the use of a mouse is inconvenient or impossible.

We designed and implemented novel gaze & key click alternatives combining eye gaze tracking and keyboard input (Letter Assignment and Offset Menu), allowing users to click targets on the screen without the mouse. These click alternatives are able to mitigate some of the inaccuracies of eye gaze trackers and the eye, resulting in an improved accuracy when compared to a naive click alternative based on direct gaze pointing and a physical button (Gaze & Click). They are still significantly slower and less accurate than the mouse, however, we believe that with more work they could become realistic mouse replacements for certain situations.

One major issue found during the experiments was calibration: it was frequently necessary to recalibrate the gaze tracker, and many participants found this tiring and time-consuming. As a consequence, the use of methods for automatic or simplified calibration should be considered. Furthermore, there are problems of the proposed click alternatives that should be addressed, e.g. the assignment of unintuitive letters to targets in Letter Assignment and difficulties with the de-selection of potential targets in Offset Menu.

Finally, there are some open questions. For example, in how far did touch typing skills affect the performance of Letter Assignment? What are the long-term effects of using gaze & key click alternatives? How do such click alternatives compare to other pointing devices such as trackpads and touchscreens? We hope to answer some of these questions in future work.

Acknowledgements. We would like to acknowledge Abdul Moiz Penkar for his work on the web browser implementation, and all our participants without whom this study would not have been possible.

References

1. Porta, M., Ravelli, A.: WeyeB, an Eye-Controlled Web Browser for Hands-Free Navigation. In: Proceedings of the Conference on Human System Interactions (HSI), pp. 210–215. IEEE (2009)
2. Porta, M., Ravarelli, A., Spagnoli, G.: ceCursor, a Contextual Eye Cursor for General Pointing in Windows Environments. In: Proceedings of the Symposium on Eye-Tracking Research & Applications, pp. 331–337. ACM (2010)
3. Skovsgaard, H.: Noise Challenges in Monomodal Gaze Interaction. PhD thesis. Köpenhamn: IT University of Copenhagen (2012)
4. Bates, R., Istance, H.: Zooming Interfaces!: Enhancing the Performance of Eye Controlled Pointing Devices. In: Proceedings of the Fifth International ACM Conference on Assistive Technologies, pp. 119–126. ACM (2002)

5. Penkar, A. M., Lutteroth, C., Weber, G.: Designing for the Eye: Design Parameters for Dwell in Gaze Interaction. In: Proceedings of the Australian Computer-Human Interaction Conference (OzCHI), pp. 479-488. ACM (2012)
6. Wandell, B. A.: Foundations of Vision. Sinauer Associates (1995)
7. Zhang, X., Ren, X., Zha, H.: Improving Eye Cursor's Stability for Eye Pointing Tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 525–534. ACM (2008)
8. Sibert, L.E., Jacob, R.J.K.: Evaluation of Eye Gaze Interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 281–288. ACM (2000)
9. Ohno, T.: Features of Eye Gaze Interface for Selection Tasks. In: Proceedings of the Asia Pacific Conference on Computer Human Interaction (APCHI), pp. 176–181. IEEE (1998)
10. Urbina, M.H., Lorenz, M., Huckauf, A.: Pies with Eyes: the Limits of Hierarchical Pie Menus in Gaze Control. In: Proceedings of the Symposium on Eye-Tracking Research & Applications (ETRA), pp. 93–96. ACM (2010)
11. Ware, C., Mikaelian, H.H.: An Evaluation of an Eye Tracker as a Device for Computer Input. In: Proceedings of the SIGCHI/GI conference on Human Factors in Computing Systems and Graphics Interface (CHI), pp. 183–188. ACM (1987)
12. Kumar, M., Klingner, J., Puranik, R., Winograd, T., Paepcke, A. Improving the accuracy of gaze input for interaction. In: Proceedings of the Symposium on Eye Tracking Research & Applications, pp. 65-68. ACM (2008)
13. Kumar, M., Paepcke, A., Winograd, T.: EyePoint: Practical Pointing and Selection using Gaze and Keyboard. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 421-430. ACM (2007)
14. Ashmore, M., Duchowski, A.T., Shoemaker, G.: Efficient Eye Pointing with a Fisheye Lens. In: Proceedings of Graphics Interface (GI), pp. 203–210. Canadian Human-Computer Communications Society (2005)
15. Käser, D.P., Agrawala, M., Pauly, M.: Fingerglass: Efficient Multiscale Interaction on Multitouch Screens. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1601–1610. ACM (2011)
16. Chin, C.A., Barreto, A., Cremades, J.G., Adjouadi, M.: Integrated Electromyogram and Eye-Gaze Tracking Cursor Control System for Computer Users with Motor Disabilities. Journal of Rehabilitation Research and Development 45(1), 161-174 (2008)
17. Zhai, S., Morimoto, C., Ihde, S.: Manual and Gaze Input Cascaded (MAGIC) Pointing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 246–253. ACM (1999)
18. Blanch, R., Ortega, M.: Rake Cursor: Improving Pointing Performance with Concurrent Input Channels. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1415–1418. ACM (2009)
19. Kumar, M., Winograd, T.: GUIDe: Gaze-Enhanced UI Design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - Extended Abstracts (CHI EA), pp. 1977–1982. ACM (2007)
20. Penkar, A. M., Lutteroth, C., Weber, G.: Eyes Only: Navigating Hypertext with Gaze. In Proceedings of the IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT), pp. 153-169. Springer (2013)
21. Brooke, J.: SUS: A 'Quick and Dirty' Usability Scale. In: Jordan, P. W., Thomas, B., Weerdmeester, B. A., McClelland, A. L. (eds.), Usability Evaluation in Industry, pp. 189-194. Taylor & Francis (1996)