

*Citation for published version:*

Ramirez Gomez, A, Clarke, C, Sidenmark, L & Gellersen, H 2021, Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye. in *Proceedings ETRA 2021 Full Papers - ACM Symposium on Eye Tracking Research and Applications*. Association for Computing Machinery, ACM Symposium on Eye Tracking Research & Applications, 24/05/21.

*Publication date:*  
2021

*Document Version*  
Peer reviewed version

[Link to publication](#)

## University of Bath

### Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

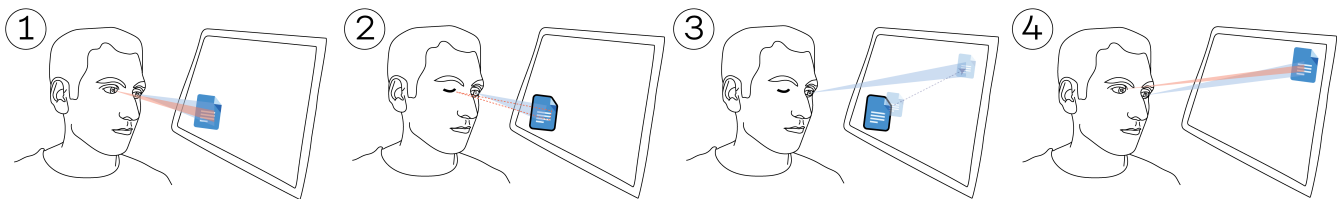
# Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye

Argenis Ramirez Gomez  
Cardiff University  
Cardiff, United Kingdom  
ramirezgomeza@cardiff.ac.uk

Ludwig Sidenmark  
Lancaster University  
Lancaster, United Kingdom  
l.sidenmark@lancaster.ac.uk

Christopher Clarke  
University of Bath  
Bath, United Kingdom  
cjc234@bath.ac.uk

Hans Gellersen  
Aarhus University  
Aarhus, Denmark  
hwg@cs.au.dk



**Figure 1: Gaze+Hold uses explicit closing of one eye to modulate gaze input from the open eye, demonstrated here with drag and drop. (1) The user looks at the interface without triggering any effects; (2) on left eye closure, the object is selected; (3) dragging is enabled via continuous gaze input from the open eye; (4) the interaction stops and the object is dropped when opening the left eye.**

## ABSTRACT

The eyes are coupled in their gaze function and therefore usually treated as a single input channel, limiting the range of interactions. However, people are able to open and close one eye while still gazing with the other. We introduce Gaze+Hold as an eyes-only technique that builds on this ability to leverage the eyes as separate input channels, with one eye modulating the state of interaction while the other provides continuous input. Gaze+Hold enables direct manipulation beyond pointing which we explore through the design of Gaze+Hold techniques for a range of user interface tasks. In a user study, we evaluated performance, usability and user's spontaneous choice of eye for modulation of input. The results show that users are effective with Gaze+Hold. The choice of dominant versus non-dominant eye had no effect on performance, perceived usability and workload. This is significant for the utility of Gaze+Hold as it affords flexibility for mapping of either eye in different configurations.

## CCS CONCEPTS

• Human-centered computing → Interaction techniques.

## KEYWORDS

Eye Tracking, Gaze Interaction, Winks, Gaze Pointing, Closing eyelids, Direct Manipulation, Design

## ACM Reference Format:

Argenis Ramirez Gomez, Christopher Clarke, Ludwig Sidenmark, and Hans Gellersen. 2021. Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye. In *2021 Symposium on Eye Tracking Research and Applications (ETRA '21 Full Papers)*, May 25–27, 2021, Virtual Event, Germany. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3448017.3457381>

## 1 INTRODUCTION

Eyes-only control is compelling for a wide range of contexts where other input is unavailable, not as ready at hand, or inconvenient. Eye movement is fast and requires less energy and effort than input with head or hands [Land and Tatler 2009; Sidenmark and Gellersen 2019], and interfaces that can be operated with just the eyes have for instance been developed for accessibility [Grauman et al. 2003; Hutchinson et al. 1989; Wobbrock et al. 2008], instant control [Esteves et al. 2015; Velloso et al. 2016] and discreet input [De Luca et al. 2007; Khamis et al. 2017].

Existing eyes-only interfaces are based on continuous gaze tracking for pointing, complemented by primitives for discrete

selection, including dwell time [Jacob 1990; Ware and Mikaelian 1986], long blinks [Grauman et al. 2003], saccadic gestures [Drewes and Schmidt 2007; Heikkilä and Rähä 2012; Wobbrock et al. 2008], and smooth pursuit [Velloso et al. 2017; Vidal et al. 2013]. This has two key limitations. First, the pointer moves with every eye movement, with no distinction of whether the movement was intended as pointing action, incidental to viewing, or caused by a visual distraction. This creates problems of feedback (whether to display a cursor [Porta et al. 2010]) and unintended input (Midas touch [Istance et al. 2008; Jacob 1990]). Secondly, use of the eyes is exclusive for either pointing or click-equivalent action, unlike with conventional pointing where positional input can concur with discrete input (e.g., mouse movement while pressing a button). As a consequence, continuous direct manipulations (e.g., moving an object) are not well supported, as they have to be mapped to successive interaction steps (selection of command, object and target position).

In this work, we introduce *Gaze+Hold* to extend expressiveness of eyes-only interaction. The key idea is to leverage the eyes as two separate sources of input that can be combined to support continuous direct manipulation. Even though we have two eyes, they are commonly treated as a single input channel. Eye trackers fuse input from both eyes to provide a continuous stream of gaze estimates, as the eyes are coupled in their gaze function. However, humans can voluntarily open and close each eye independently, and keep either eye closed while gazing with the other [Jota and Wigdor 2015]. The eyes can therefore be considered as separate input channels, where the open/close state of either eye can be used to modulate gaze input observed from the open eye. *Gaze+Hold* leverages this for interactions that are initiated by closing of an eye, performed with continuous gaze, and completed when the eye is opened. Figure 1 exemplifies use of the technique to manage Midas Touch and support drag-and-drop interaction.

The temporary switching of a pointer by discrete input is pervasive in state of the art interfaces for dragging, panning, zooming and other forms of continuous manipulation, and *Gaze+Hold* opens up a design space for such tasks to be performed with just the eyes. The eyes are distinct in that either eye can be used to gaze and point while the other is closed, although most people have a dominant eye [Elbaum et al. 2017]. This leads to specific design considerations for *Gaze+Hold* that we develop through exploration of range of novel eyes-only manipulation techniques for common continuous input tasks.

*Gaze+Hold* relies on one-eyed gaze as continuous input, raising questions of usability and performance that we assessed in a user study with 12 participants. With a binocular eye-tracker, gaze estimates are expected to be less accurate when they are temporarily based on a single eye, while coordination of gaze with eye closing/opening increases effort. Users are nonetheless as effective with *Gaze+Hold* for drag-and-drop tasks as with a baseline of binocular gaze and button click. Speed and accuracy are lower than with both eyes open, but users gain the advantage of relying solely on their eyes for input. Notably, all users were able to use *Gaze+Hold* with either eye closed, and the choice of dominant versus non-dominant eye for manipulation had no significant effect on performance, usability and perceived workload.

## 2 RELATED WORK

Jacob's work thirty years ago was seminal in considering eye movement from a design perspective of supporting interaction with graphical user interfaces [Jacob 1990]. Eyes-only computer control had been demonstrated before [Hutchinson et al. 1989; Shaw et al. 1990] but Jacob's work established a general style of eye movement for discrete selection of objects and commands by pointing, with dwell time as a click method. However, the work also discussed that manipulation tasks such as moving an object require complementary input akin to holding down a mouse button, in order to switch between gaze pointing and manipulation modes. Jacob and others have addressed this by multimodal combination of gaze with key or button input [Hild et al. 2016; Jacob 1990; Kumar et al. 2007], and in related work gaze has been complemented by touch [Pfeuffer et al. 2014; Stellmach and Dachselt 2013; Turner et al. 2015, 2014] and mid-air gestures [Pfeuffer et al. 2017] in similar fashion. This work, in contrast, contributes an eyes-only approach, based on the idea of one eye controlling discrete state while the other points.

Historically, much of the work on eyes-only interfaces is framed as mouse emulation [Lankford 2000]. Blinks and winks have been explored to mimic button events in order to make techniques designed for the mouse accessible [Kowalczyk and Sawicki 2019; Missimer and Betke 2010], whereas the focus of this work is on design for eyes-only direct manipulation. In analogy to mouse behaviour, Istance and colleagues discussed different modes for gaze input, including discrete selection (point and click), continuous manipulation (dragging) and disengagement from cursor movement and control to let the eyes freely look at the interface without unintended input [Istance et al. 2008]. In their work on Snap Clutch, the modes are activated by quickly glancing off-screen (up, down, left or right). *Gaze+Hold* is conceptually different as it supports a default state from which specific input modes are activated by closing an eye, and active only for the duration of eye closure. This lets users keep their gaze focus on the object they intend to manipulate, with default gaze behaviour instantly resumed once an input task is completed by opening the closed eye.

The key idea underpinning our work is to use closing and opening of one eye, to modulate gaze input performed with the other. The literature refers to the closing/opening of one eye as wink, contrasting blinks performed with both eyes [Kaufman et al. 1993; Kowalczyk and Sawicki 2019; Missimer and Betke 2010; Singh and Singh 2018]. A wink is a holistic gesture whereas we build specifically on eye-closing and eye-opening events to trigger and delimit direct manipulation. However, there is a large body of work on which we build. Some of the earliest eye control interfaces were based entirely on blinks and winks, highlighting the utility of eye closure as input [Shaw et al. 1990]. Single-eye closure has the advantage that it rarely occurs involuntarily and is less prone to false positive input [Kowalczyk and Sawicki 2019]. Most people can use either eye to wink but some tend to close both eyes when attempting to wink with the dominant eye [Palmer 1976]. Importantly for this work, movement of the eyelid is independent of movement of the eyeball [Jota and Wigdor 2015], and people still have full gaze control when they close one eye. In manual targeting and manipulation tasks, closing of one eye tends to only minimally effect performance, with performance slightly better when the dominant as opposed

to non-dominant eye guides the task [Coren 1999; Marotta et al. 1995]. For our work, this prompts comparison of dominant versus non-dominant eye for “one-eyed pointing”.

Prior work has also considered design of interaction techniques based on eye closure. Winks and blinks have been widely considered as a delimiter for gaze input [Grauman et al. 2003; Missimer and Betke 2010] as well as other modalities [Wang and Grossman 2020]. It has also been shown that winks and blinks can be effective as command gesture, for example as shortcut to smartphone functions [Ku et al. 2018], for switching between different views [Hemmert et al. 2008; Jota and Wigdor 2015], or as part of larger eye gesture vocabularies for users with severe motor impairment [Zhang et al. 2017]. There is also work on input with both eyes closed, for example to perform gaze gestures for password entry [Findling et al. 2019] or as a game mechanic [Ramirez Gomez and Gellersen 2019a, 2020]. Jota and Wigdor provided a wider exploration of the design space of eyelid gestures, noting for example that duration of eye closure can also be used for input [Jota and Wigdor 2015]. Hemmert and colleagues provided compelling use cases for closure of one eye to temporarily change the perspective of the user, for instance to zoom into a scene to see detail, or zoom out to a macro-view from a different view can be selected by the mouse [Hemmert et al. 2008]. Gaze+Hold is likewise based on eye closure to control state, but extends beyond view control to modulation of input and eyes-only direct manipulation.

Diverse forms of gaze interaction exist beyond point and dwell, including saccadic gestures (e.g. left, right, up, down) [Heikkilä and Riihinen 2012; Zhang et al. 2017], glances off-screen [Istance et al. 2008] or to confirmation targets [Lutteroth et al. 2015], gestures “drawn” with multiple saccades [Bâce et al. 2016; Drewes and Schmidt 2007; Wobbrock et al. 2008], and smooth pursuit eye movement in correlation with moving stimuli [Esteves et al. 2015; Gomez and Gellersen 2018; Sidenmark et al. 2020; Velloso et al. 2017; Vidal et al. 2013]. In principle, any of these can be combined with Gaze+Hold, as they constitute gaze input that can be performed with one eye. However, closing one eye also presents an unusual form of input and constraint to vision and gaze. In this respect, Gaze+Hold relates to work exploring gaze under constraints, for example avoiding direct gaze and relying on peripheral vision [Ramirez Gomez and Gellersen 2019b], and forms of gaze interaction beyond conventional looking [Ramirez Gomez and Gellersen 2019a].

### 3 GAZE+HOLD CONCEPT AND INTERACTION TECHNIQUES

Gaze+Hold defines a distinct style of eyes-only interaction. At a higher level, this is characterised by support of different states for gaze interaction that are controlled by eye closure, and at technique level by continuous gaze input with one eye, delimited by closing and opening of the other. Gaze+Hold interactions are performed within specific input modes associated with left or right eye, and defined by:

- (1) Closing of one eye initiates input;
- (2) Movement of the other eye provides continuous input;
- (3) Opening of the eye to complete the interaction.

We explore Gaze+Hold through the design of concrete interaction techniques for general user interface tasks. All techniques and application examples were developed using Unity Game Engine in C# in 2D with both custom and creative commons graphics. We used the Tobii Gaming SDK for Unity to get the position of the gaze point and the Tobii EyeX Framework to separately detect eye-closing and eye-opening events, based on absence versus presence of left or right eye from visual tracking. The system automatically filters involuntary blinks, during which both eyes become absent from tracking for a short duration (~250ms).

#### 3.1 Selecting and Moving Objects

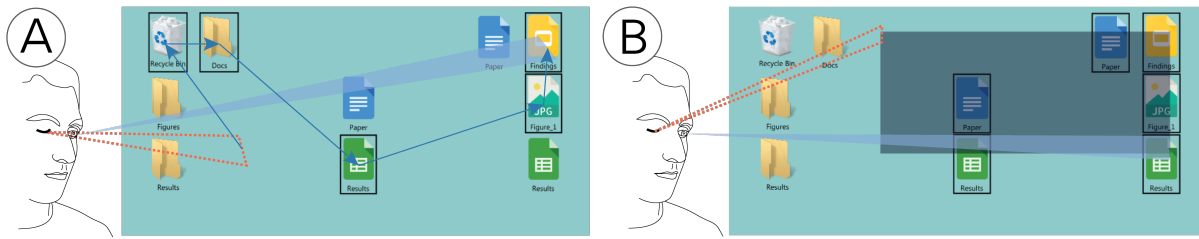
Selection and moving of objects are canonical tasks in graphical user interfaces. We implemented a set of Gaze+Hold techniques equivalent to the most basic mouse operations. In Gaze+Hold, a selection is performed by gazing at an object, and closing and opening of one eye while gaze remains in position. This is effectively point and click with a wink gesture as click method [Missimer and Betke 2010]. However, selection can also be implemented as two-step with eye-closing to hover, gaze to evaluate feedback, and eye-opening to confirm.

Drag-and-drop is the simplest way of moving an object in an interface. With Gaze+Hold, an object can be moved by gazing at it, eye-closing to select it, gaze pointing to a target location, and eye-opening to drop the object (Figure 1). Like selection, this is a basic technique where gaze input is purely positional. It can be performed quickly with a direct gaze shift to the target, but also allows for multiple gaze shifts, for the eye to search and evaluate target locations before committing to drop. As a variant, copy-paste can be supported in the same way, dragging a copy rather than the object.

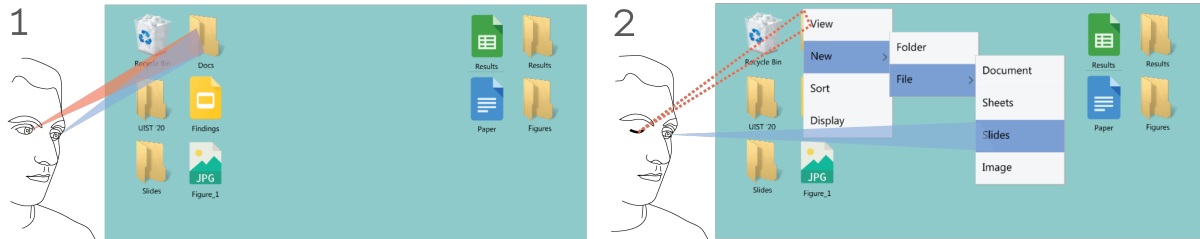
#### 3.2 Multiple Object Selection

Closure of one eye can be used to frame selection of multiple objects. Once activated by eye-closure, this technique selects all objects crossed by gaze during eye closure (Fig. 2, A). The technique is not position-dependent; it can be initiated on the first object to select, but also in any neutral part of the interface. Input is positional and can be fast as users do not need to fixate on individual objects, for instance selecting a row of objects by gazing across. However, crossing can be prone to unintended selection of other objects (Midas Touch). The technique could be varied to mitigate for this by using dwell time or brief blink (closure also of the other eye) for selections during eye closure.

Alternatively, multiselection can be performed by a selection box, as also common with mouse interaction. The gaze position at eye-closing defines one corner of the box. Gaze during eye closure controls the opposite corner of the box, such that it can be pulled out over the set of objects intended for selection, finally confirmed by eye-opening (Fig. 2, B). Gaze input is positional and can be completed with a single gaze saccade from the start to the endpoint. This can also be applied for text selection, for example by closing an eye at the start of a passage and opening it at the end.



**Figure 2: Selection of multiple objects. A) *Multiselection* is initiated by eye-closure and selects all objects crossed by gaze, dwelled or blinked at until eye-opening. B) *Multiselection* is performed by drawing a selection box.**



**Figure 3: Context Menu. (1) On closure of one eye, a menu opens and (2) can be freely explored and navigated by gaze, with final selection of an item committed by eye-opening.**

### 3.3 Menu Navigation and Selection

Design for eyes-only menu interaction is challenging as it requires mechanisms for menu opening, navigation, and item selection. This involves specific Midas Touch issues as users may dwell longer on specific items when they consider choices, commonly addressed with longer dwell times or additional interaction steps [Figueiredo and Fonseca 2018; Jacob 1990; Špakov and Miniotas 2005]. Gaze+Hold provides an alternative, illustrated with a context menu (Fig. 3). The menu is opened by eye-closing and allows items to be visually explored with hover feedback. If an item has a sub-menu associated, this also opens in response to hover, but closes when gaze moves to another item at parent-level. A selection is completed by eye-opening on an item, but selection can also be abandoned by gazing off-menu prior to eye-opening. The design illustrates how separate input from the eyes is utilised to support visual exploration while avoiding Midas Touch. One eye looks and points without risk of unintended selection, while the other performs discrete input to trigger a menu, and to confirm a selection.

### 3.4 Zooming or Scaling

Gaze+Hold can be used to control zoom or scaling of an object or interface. Initiated by eye-closing, the focus of the zoom is fixed to the initial gaze point. The continuous input of the open eye can be used to zoom in or out at a fixed speed by performing discrete gaze gestures relative to the focus point (looking up/downwards) and guided by visual indicators (Fig. 4, A). The technique is robust as it relies only on direction of gaze movement up or down, but the design forces the user to direct their gaze away from an initially selected focus. As an alternative, separate states can be mapped to left and right eye (one to zoom in, the other to zoom out), enabling

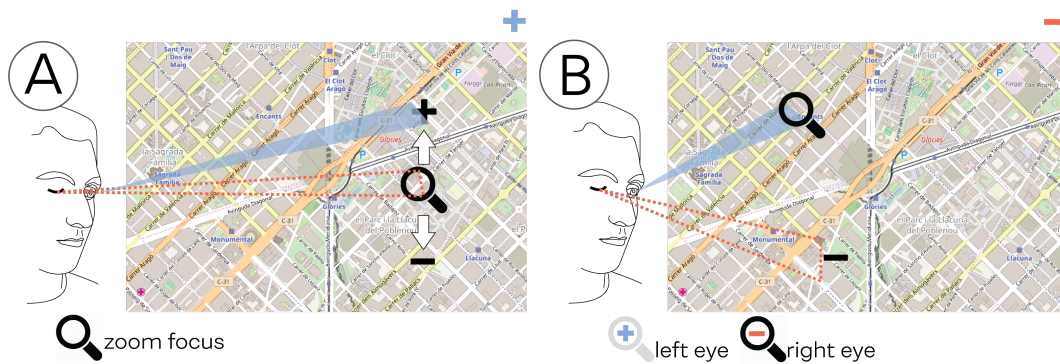
gaze during eye closure to be used to continually control the focus point (Fig. 4, B). Such a mapping makes it possible, for example, to adjust the focus based on detail that is gradually revealed by zoom.

### 3.5 Panning and Scrolling

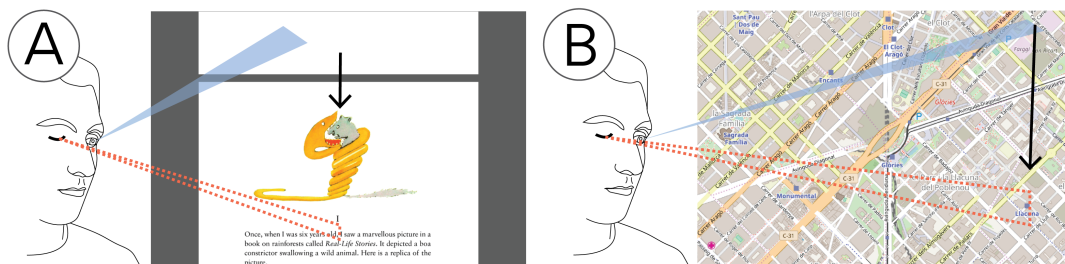
Panning and scrolling are operations to adjust the viewport over a larger space or document. Gaze lends itself to shifting the view over shorter distances where gaze is natural in selecting content to which the view should be adapted. After triggering by eye-closure, the canvas is shifted in the direction from the gaze position to the display centre, which is marked as reference point (Fig. 5, A). The mapping has a self-centering effect inspired by *GazeHorizon* [Zhang et al. 2014]: as users maintain their gaze on content of interest during panning, the viewport will become centred around it. Gaze can also be shifted to content gradually revealed, to pan further, and eye-opening completes the interaction. The gaze input is positional but used relative to the display centre. In an alternative mapping, the gaze position at eye-closing can be used to set an anchor point in the viewport. Gaze can then be moved to select content that is panned or scrolled toward the anchor position (Fig. 5, B). Such a mapping affords control over adjustment of the view, for example to align particular content with one of the edges of the viewport.

### 3.6 Navigation and Teleportation

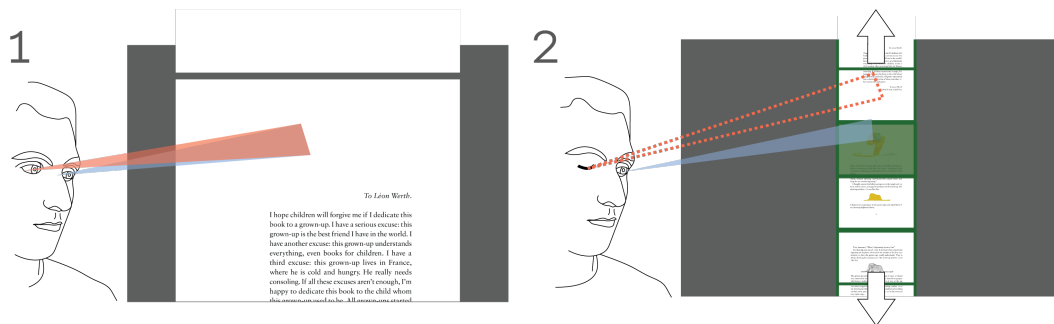
Panning and scrolling with Gaze+Hold is effective for view adjustment over shorter distance but less so for navigation of larger spaces or documents, as holding an eye closed for longer may induce fatigue. However, eye closure can also be mapped to provide users with a different view or perspective, as explored e.g. in *sniper zoom* [Hemmert et al. 2008]. A quick change in view can be useful



**Figure 4: Zooming with Gaze+Hold.** A) uses gaze gestures up or down, to zoom in/out from a focus point at which the technique was invoked; B) uses closing of one eye to zoom in, and closing of the other to zoom out, while the eye that remains open determines the focus point.



**Figure 5: Scrolling/Panning.** The document/map is shifted in the direction from gaze position to (A) the centre of the screen; (B) the anchor point defined by the gaze position at eye-closing.



**Figure 6: Navigation and Teleportation.** (1) On eye-closing, Gaze+Hold zooms out the document so (2) the user can look at a position in the document to navigate to when opening the closed eye (Navigate-Select).

for cross-referencing information from another document or screen when inputting data, or temporarily overlaying additional information, e.g. textures in 3D modeling. Closing of an eye triggers a change in view, to an overview of the interaction space, such as a zoomed out document. The overview can be navigated by gaze to select a different part of the document, at which detailed viewing is resumed after completion of the interaction by eye-opening (Fig. 6). The navigation is over discrete parts of the interface, and lends

itself to moving for example between pages in a document, images in a gallery, tabs in a web browser, or applications in the operating system. Alternatively, eye-closing can trigger a “world in miniature” view of the interaction space, and gaze can manipulate the new view with a moving window to select the part of the world that will be in view after completion of the navigation on eye-opening. This lends itself to navigation of interactive worlds, and for example teleportation in games.



### 3.7 Line Drawing and Measuring Distances

Gaze is efficient for selecting lines on the interface as start and end point of a saccade. With Gaze+Hold, eye-closing can initiate line selection and eye-opening complete it, for line drawing or measuring distances between points in the interface. More complex paths can be created using anchor points, for example with brief closure of the open eye at intermittent points, or by using dwell time. However, drawing of more complex paths may cause fatigue as it requires closure of one eye for prolonged periods.

## 4 EVALUATION OF PERFORMANCE AND USABILITY

The design exploration shows that Gaze+Hold enables a wide range of direct manipulation tasks to be performed with input solely from the eyes. However, the techniques demonstrated depend on accurate gaze input while one eye is closed. This raises usability questions from both a technical and user perspective, as it depends on the user's ability to keep one eye closed while using the other for input and the system's performance in estimating gaze when one eye is temporarily unavailable for tracking. Due to the requirement of the user to keep one eye shut for the duration of input, this is succinctly different to the conventionally defined monocular accuracy, in which only one eye is being tracked by an eye tracker yet the other eye is open [Gibaldi et al. 2017]. The primary goals for evaluation were therefore to establish practicality and effectiveness of Gaze+Hold, and to compare performance against a baseline of gaze input with both eyes open and manual toggling of input state.

Previous work has suggested that users are more accurate when fixating with their dominant eye when the other is covered [Nyström et al. 2013; Simonsz and Bour 1991]. This raises the question of how choice of eye (dominant/non-dominant) affects the performance, workload and usability of Gaze+Hold. A related question pertains to preference for using one eye over the other, and how this is manifest in choices users make when interact spontaneously. This is an important factor when considering how the technique could be incorporated and configured for different applications: does eye dominance have a similar impact on user preference to (e.g.) hand dominance? We designed our evaluation therefore to gain insight specifically into use of dominant versus non-dominant eye for Gaze+Hold interaction.

### 4.1 Participants and Apparatus

Twelve participants took part in the study (8 male, 4 female, age:  $M=25$ ,  $SD=3$ ). Four of the participants wore glasses, two contact lenses, and six had uncorrected vision. Four were experienced with eye tracking and gaze interaction. Ten had right-eye dominance, and two had left-eye dominance according to the eye-dominance test [Vision 2017]. All participants ability to wink was verified prior to the study. There were no study participants who were unable to wink. As apparatus, we used a Tobii EyeX eye tracker under a 27" monitor (Resolution: 1920x1080; Aspect Ratio: 16:9) located 40cm from the user. The study application uses the same implementation pipeline as described above (section 3) for the implementation of techniques.

### 4.2 Design and Procedure

Upon arrival, participants were presented with an information sheet and consent form, and filled in a demographics questionnaire. They were then asked to wink at the researcher to note their spontaneous eye choice for winking, and verify they could wink. Finally, the eye tracker was calibrated for each participant using Tobii's standard five-point procedure. The study then proceeded with two parts, a comparative evaluation followed by an evaluation of spontaneous choice of dominant versus non-dominant eye for Gaze+Hold interaction.

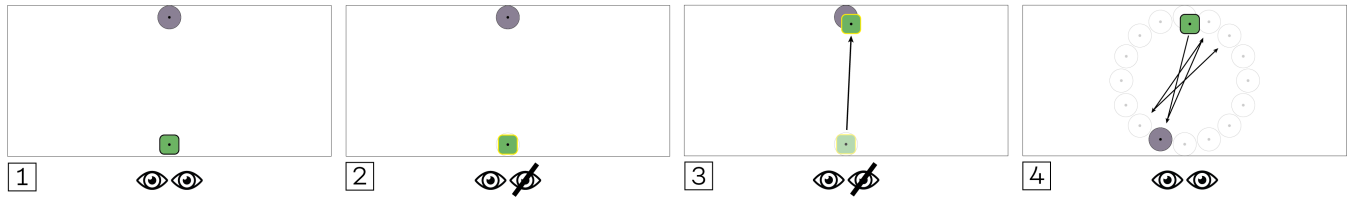
**4.2.1 Comparative evaluation.** The first part of the study was a factorial experiment with three conditions (*Dom*, *NonD* and *Ref*) comparing performance of drag-and-drop with the dominant eye, non-dominant eye and both eyes. In *Dom*, participants used Gaze+Hold with the dominant eye for continuous input while closing the non-dominant eye, and in *NonD* with the opposite mapping. *Ref* represented a reference condition in which the task was performed with both eyes open, combined with key press/release to manually initiate and complete the interaction. The combination of gaze with a manual trigger has been shown to be comparable with the mouse in throughput [Zhang and MacKenzie 2007] and was adopted as baseline to assess trade-off in performance for gaining the advantage of eyes-only control.

Drag-and-drop was used as a canonical user interface task for the study, representative of continuous interactions enabled by Gaze+Hold. Drag-and-drop uses a continuous output position allowing for measurement of the speed and accuracy of the technique, and is a task that uses visual feedback which may affect usability. Resembling the multi-directional Fitts' Law task for evaluation of gaze pointing [San Agustin et al. 2009; Vertegaal 2008], we designed the task with 16 targets arranged in a circle around the center of the screen (500px diameter,  $\sim 19^\circ$  visual angle). Participants were tasked to drag a square-shaped object (75x75px,  $\sim 5^\circ$ ) to a target across the circle and to drop it centrally on the target. Targets were circular with 130px ( $\sim 5^\circ$ ).

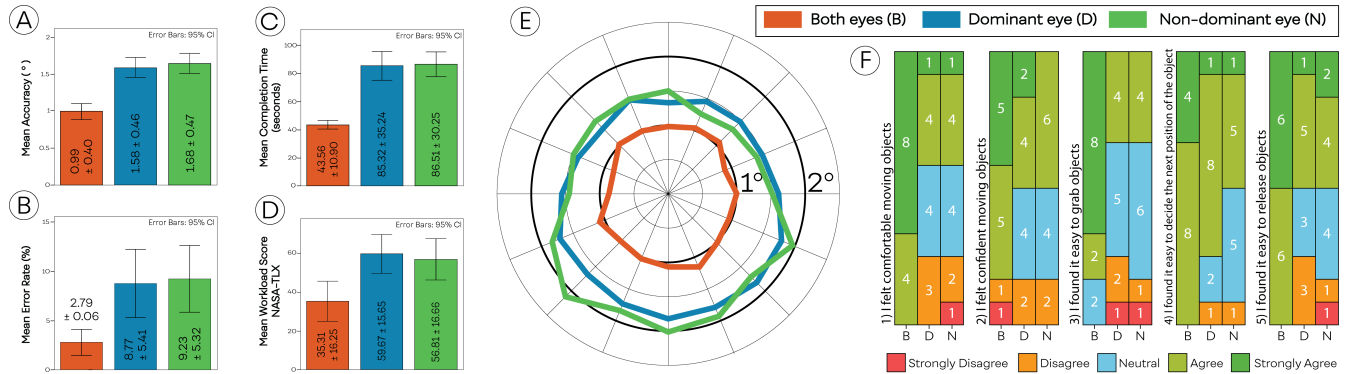
Figure 7 illustrates the performance task from the user's perspective. Initially the object and a first target are shown. Once the object has been dropped on the target, it is removed from the display and a new target is shown across the circle. The order of the targets follows the Fitts' Law task procedure. The task continued for two loops, for a total of 32 drag-and-drop interactions. The performance task was repeated for each condition, in counterbalanced order. Before the task, participants had the opportunity to practice with the techniques. Participants performed the task four times (128 target hits) in each condition, and were instructed to complete each task as quickly and as accurately as possible.

We measured the time taken for each iteration of the performance task, dropping the objects on a total of 32 targets, and the accuracy of each drop as distance of the object's centre from the target's centre. After each condition participants filled in a custom questionnaire with five 5-point Likert items to assess usability, and a NASA-TLX questionnaire to measure task workload [Hart and Staveland 1988].

**4.2.2 Spontaneous Preference.** The second part of the study was designed to elicit spontaneous choice of eye for Gaze+Hold. Users



**Figure 7: Performance Task.** The user is presented is when a green object and a target on which to drop it (1). The drag-and-drop operation is initiated by eye closure (2), performed with the open eye (3) and completed with eye opening, upon which the next target is shown (4). The task continuous around the circle of targets, until two rounds are completed.



**Figure 8: Results of the performance evaluation, comparing Gaze+Hold with dominant versus non-dominant eye, and a baseline of gaze with both eyes open in which a manual trigger was used to delimit gaze input. A) Total Accuracy and error rate; B) Error Rate; C) Task Completion Time; D) NASA-TLX Workload Score; E) Mean accuracy for each target's direction; F) Usability Questionnaire.**

were presented with a desktop application scenario, in which they had to use drag-and-drop to move files into folders, both represented by icons. A total of 16 different files (4 each: docs, sheets, slides and images) were scattered on the desktop display, and each had to be dragged and dropped onto their corresponding folder. During the task, we recorded for each interaction which eye was closed when a file object was dragged. After the task, NASA-TLX was administered to measure task workload.

### 4.3 Results

The results from the performance evaluation are summarized in Figure 8. Accuracy results are reported in visual angle. For accuracy, we only consider successful “drops” of the object on the target. Error is defined as failure to complete a trial on first attempt, i.e. when more than one drag and drop movement was required to position the object on the target. We compared the conditions using a Friedman test. Post-hoc tests were conducted using Bonferroni-corrected Wilcoxon signed-rank test.

**4.3.1 Gaze+Hold Performance.** There were statistically significant differences for accuracy ( $\chi^2(2)=440.13$ ,  $p<0.001$ ); completion time ( $\chi^2(2)=72.67$ ,  $p<0.001$ ), and error rate ( $\chi^2(2)=10.50$ ,  $p=0.005$ ), see Fig. 8, A-C.

Using post-hoc analysis, we first look at the performance comparison between dominant and non-dominant eye. There were no significant differences using Gaze+Hold with either eye for accuracy ( $Z=-1.75$ ,  $p=0.079$ ); completion time ( $Z=-0.53$ ,  $p=0.59$ ) or error ( $Z=-0.63$ ,  $p=0.53$ ). This demonstrates that the eye used for Gaze+Hold does not affect its performance. The performance of Gaze+Hold is a combination of both eye tracker accuracy, and user ability to gaze with one eye closed. As expected, we observed that the performance is affected by the use of only one eye in comparison to both eyes: *Ref* was statistically significantly more accurate (*Dom*:  $Z=-18.903$ , *NonD*:  $Z=-19.93$ ,  $p<0.001$ ); faster (*Dom*:  $Z=-21.87$ , *NonD*:  $Z=-27.09$ ,  $p<0.001$ ) and had fewer errors (*Dom*:  $Z=-2.51$ ,  $p=0.01$ , *NonD*:  $Z=-0.63$ ,  $p=0.53$ ) than Gaze+Hold with either eye.

Fig. 8, E shows the mean accuracy for each test condition during the performance task in a polar diagram representing all the direction/locations of the targets. Results show that using Gaze+Hold resulted in more accurate drag and drops when performing the drag upwards compared with dragging downwards. It is unclear why this is the case. Previous work has noted how there can be deviation in the eye movements after prolonged, voluntary blinks [Collewyn et al. 1985], and this maybe a similar phenomenon. This could also be an artefact of when the eye tracker deems a gaze reading valid, i.e. when the eye faces down, the pupil is visible earlier upon eye opening.



**Table 1: Usability results comparing Gaze+Hold with dominant eye (Dom), non-dominant eye (NonD) and the reference technique (Ref).**

Questionnaire Statement	Friedman Test	Dom - NonD	Dom - Ref	NonD - Ref
(1) "I felt comfortable moving objects"	$\chi^2(2)=16.65, p<0.001$	$Z=-0.30, p=0.763$	$Z=-3.0, p=0.003$	$Z=-2.86, p=0.004$
(2) "I felt confident moving objects"	$\chi^2(2)=6.53, p=0.038$	$Z=-1.03, p=0.305$	$Z=-1.40, p=0.161$	$p=-2.31, p=0.021$
(3) "I found it easy to grab objects"	$\chi^2(2)=17.06, p<0.001$	$Z=-0.58, p=0.564$	$Z=-2.85, p=0.004$	$Z=-2.70, p=0.007$
(4) "I found it easy to decide the next position of the object"	$\chi^2(2)=8.97, p=0.011$	$Z=-1.00, p=0.317$	$Z=-2.07, p=0.038$	$Z=-2.49, p=0.013$
(5) "I found it easy to release objects"	$\chi^2(2)=9.39, p=0.009$	$Z=-0.33, p=0.739$	$Z=-2.51, p=0.012$	$Z=-2.36, p=0.018$

**4.3.2 Usability and Workload.** Figure 8, F shows how participants rated the five usability statements. Table 1 shows that Friedman tests revealed a significant difference for all usability statements. However, as was the case with performance, there were no significant differences between the usability of using Gaze+Hold with either the dominant or non-dominant eyes. In contrast, participants found the reference technique more comfortable when moving the objects, and felt it was easier to grab the objects to begin with compared with Gaze+Hold when using either the dominant or non-dominant eye. The reference technique was also reported to be easier to use when deciding the next position compared with the non-dominant eye, and easier to release the object compared with the dominant eye.

We found a statistically significant difference on the NASA-TLX workload for the performance task ( $\chi^2(2)=15.50, p<0.001$ ). As shown in Figure 8, the task load was perceived equal between the two Gaze+Hold conditions ( $Z=-0.82, p=0.409$ ) but Ref was perceived significantly different than Dom ( $Z=-3.06, p=0.002$ ) and NonD ( $Z=-2.98, p=0.003$ ). The mental demand, physical demand, and effort dimensions were the main contributors to the workload. The workload for the desktop task was  $51.64 \pm 15.81$ , with effort, performance and mental demand as the main contributors.

**4.3.3 User Preferences.** At the start of the study, eleven participants winked at the researcher with their right eye, and one participant with their left. During the desktop files task, seven participants closed their dominant eye the most, whereas five predominantly used their non-dominant eye. Most participants' results were polarised favouring one eye versus the other one (>80%). However, three participants switched more often between each eye.

Five participants reported that they preferred their dominant eye to be closed, whereas six the non-dominant, and one was not able to decide between them. We observed a preference dissonance for four participants who reported to prefer one eye but used the other (three dominant, one non-dominant). We found a dissonance also between the spontaneous wink performed before the study and the out-loud choice (seven different) and between the wink and the eye closed during the task (five different).

## 4.4 Study Discussion

The evaluation of Gaze+Hold provides insights on the feasibility and efficacy of using the technique for eyes-only interaction, showing how participants can successfully use both eyes separately with the technique. Gibaldi et al. have shown that it is best to perform monocular calibration when measuring monocular accuracy [Gibaldi et al. 2017]. We used the out-of-the-box, standard 5-point Tobii EyeX binocular calibration which may have contributed to the decreased

accuracy we observed for Gaze+Hold. As Gibaldi et al., we observed no significant difference in accuracy between dominant and non-dominant eyes when measuring monocular accuracy. Moreover, we found no statistical differences between the eyes for performance, usability, or perceived workload for Gaze+Hold interaction. The observed dissonance between the participants' preferences further suggests they were unaware of which eye they used spontaneously with Gaze+Hold. This indicates that the eyes can be used interchangeably for interaction, providing flexibility for different configurations, and validating usability of techniques that leverage both eyes (e.g., one eye to invoke actions such as selection or drag and drop, the other to open and interact with a menu, analogue to the mouse).

The results indicate that the performance of Gaze+Hold is reduced compared with the multi-modal reference technique of gaze with manual trigger. In addition to the binocular eye tracker being calibrated for both eyes, the physiological nature of having to keep one eye closed whilst pointing with the other may be a contributing factor. We also note that the study involved interaction using Gaze+Hold at high intensity for an extended period, whereas realistic applications may involve shorter and more sporadic use of the technique. Despite this, participants were effective with the technique with either eye closed.

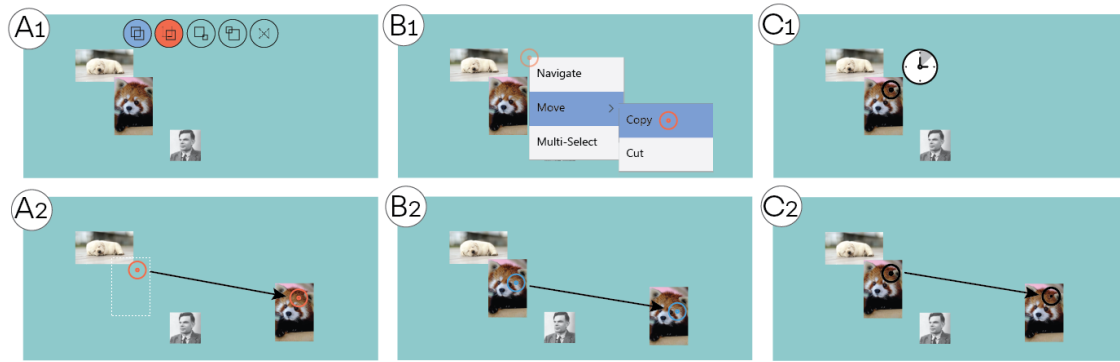
## 5 DESIGN OF GAZE+HOLD INTERACTIONS

Gaze+Hold opens up a design space for eyes-only interaction, for which we distil key design consideration based on insights from the design exploration and user study. These concern the design of techniques, as well as their integration and dynamic configuration in interfaces that scale for eyes-only access to diverse functionality.

### 5.1 Technique-level Design Considerations

**One- versus Two-State.** Techniques can have a single top-level state mapped to either eye, or two states with left- and right-eye mappings. Two-state mappings lend themselves to support complementary or opposing actions (e.g. increase/decrease). One-state mappings provide for more flexibility, for instance for user choice of preferred eye, or configuration of the other eye for other functionality to be equally "ready to hand".

**Positional versus Compound Input.** Gaze input performed during eye closure can be purely positional for pointing, where eye-closing selects a start point, and input is continually tracked for selection of an endpoint by eye-opening. Alternatively, input may involve distinct events extracted from the eye (fixations, dwell time, or any form of eye gesture) for compound input. Examples are combination of point and dwell for selection, and multi-step navigation of



**Figure 9: Gaze+Hold variants for configuration: (A) Using a Toolbar to dynamically set the action each eye closing performs (left=blue, right=red) to cut-paste with the right eye; (B) The right eye controls a pop-up Menu to assign copy-paste to the left eye; (C) Combining Gaze Mechanics: A dwell on an image after eye closing to copy-and-paste it.**

menus. Purely positional input supports fast interaction for basic tasks (e.g., moving an object) whereas compound input enables richer interaction and is unbounded with regards to forms of eye expression that can be leveraged.

*Input Granularity.* Positional input can be leveraged at different levels of granularity, fine-grained for selection of points, or coarse-grained for selection of regions. Eye-tracking has well known accuracy limitations that need to be considered in techniques that are based on fine-grained input, while techniques can be designed to operate on coarse-grained input to be more robust against gaze tracking error.

*Position-dependence.* Techniques may depend on the position at which they are invoked by closing of an eye, for example if this is used to select an object for manipulation, but can also be designed to be invoked anywhere on the interface, for example to switch into a different state.

*Absolute versus Relative Input.* Gaze input can be absolute in terms of positions in the interface, or relative in terms of offset from an anchor point. This can be a reference point in the interface, for example the centre of the display, or the gaze position at the start of the interaction (if the interaction is position-dependent).

*Discrete versus Continuous Feedback.* During the interaction, feedback may be continuous, for example with a gaze cursor or gaze trail; or discrete, for instance highlighting objects crossed by gaze, or providing feedback only at the endpoint of the interaction.

## 5.2 Interface-level Design considerations

*Interface states.* Gaze+Hold is designed to facilitate gaze input dependent on state. Gaze with both eyes open defines a default state, and gaze with left or right eye closed each define a specific input state. The direct association of different states with eye closure allows for fast switching between gaze behaviours while avoiding a moded approach. There is no presumption on the type of gaze behaviour in the different states, and in any of the states the behaviour can additionally depend on object, location or region of gaze, or other context.

*Default Behaviour.* Gaze+Hold lends itself to managing Midas Touch, by separating default behaviour from explicitly triggered

input states. A default state could be chosen accordingly, for example using gaze only for looking, only for passive feedback, or only for conservative forms of input that are not prone to unintended activation.

*Fixed versus Dynamic Configuration of Input States.* Input states associated with left and right can be statically or dynamically configured. Where an application requires only one or two forms of manipulation, these can be mapped for direct access to either eye. If more functionality needs to be supported, the mapping can be dynamic.

*Grabbing of Tools with the Eyes.* Toolbars provide access to different modes for manipulation. A unique affordance of Gaze+Hold is possible that either eye, or both eyes, can be dynamically configured with a different tool. Figure 9, A, shows an example implementation, where the user can “grab” a tool by looking at, and closing either left, right or both eyes. This can provide fast access to techniques used in alternation (e.g. for moving and resizing objects). Alternatively, a tool can also be selected from a context menu, illustrated in Figure 9, B: one eye is assigned to open a pop-up menu and the other navigates the menu to select and bind a tool.

*Modifying Input based on Gaze.* If a task involves variants for execution, the choice may also be built into the technique and triggered by variation in the eye movement performed with the open eye. The concept is illustrated in Figure 9, C: closure of one eye initiates a drag-and-drop operation, but the semantics depend on whether the user initially fixates on the selected image or immediately saccades to the target position. The initial fixation (for a specified dwell time) is associated with taking a copy of the object to be pasted, whereas the immediate saccade results in cut and paste.

## 6 DISCUSSION

Common input modalities (e.g., mouse, touch, pen) enable seamless transition between pointing and manipulation states, such as mouse clicks versus button holds; fingers to tap versus drag and pinch; and pen to tap versus draw. This is non-trivial for eyes-only interaction, where the eyes are traditionally viewed as a single source of input. This work is novel in treating the eyes as two input channels such that input state can be modified by one eye while the other eye continues to provide gaze input. The evaluation of

Gaze+Hold shows that users are able to close and open one eye while keeping gaze with the other eye steady for selection and positioning, and that they can also perform accurate gaze input with one eye while having the other closed. Speed and accuracy are not as high as with both eyes open, but users gain the ability to transition between different modes of direct manipulation with their eyes alone, without need for any complementary modality. This also provides a new way of addressing Midas Touch, using closure of one eye as deliberate signal of intent, different from other eye-based concepts such as dwell time in that it can concur with continued gaze input.

Gaze+Hold provides a well-defined interaction primitive where closing and opening of one eye delimits the action off the other. While basic in construction, the technique is highly configurable as both eyes can be associated with different modes, and as actions performed with the open eye can range from pointing to gestures, in single or multiple steps, and coupled with different feedback strategies. A significant and unexpected insight of our work is that users perform equally well with dominant and non-dominant eye, enabling both eyes to be utilised and mapped flexibly for input. The result is a compelling design space, not only for construction of novel eyes-only interaction techniques, but also the provision of more complex eyes-only interfaces in which the eyes are dynamically configured for eyes-only access to a wider range of functionality.

It is interesting to consider how Gaze+Hold can be combined with other gaze-based techniques. Smooth pursuit, for example, can be used to select objects based on relative movement [Sidenmark et al. 2020; Vidal et al. 2013], which requires objects to be in motion. To avoid continual motion, Gaze+Hold could be used to initiate display of target motion by closure of one eye, followed by pursuit selection with the other. For selection of objects in 3D, Gaze+Hold could be extended with a recently proposed technique that exploits the vestibulo-ocular reflex (VOR) to disambiguate targets by focal distance and unlike vergence requires only one eye to be tracked relative to the head [Mardanbegi et al. 2019].

We explored Gaze+Hold for common user interface tasks with examples drawn from desktop applications. The concept, however, is more generally applicable to other contexts. In a desktop context, Gaze+Hold may be of interest primarily for accessibility whereas other applications may take advantage of the technique for eyes-only control when interfaces are beyond reach, the user's hands are busy, or when touch is undesirable (e.g., contamination-free input on public devices). Users are effective with Gaze+Hold without learning curve, rendering the technique suitable for casual and spontaneous use. However, it is unusual in relying on one-eyed gaze and this may be fatiguing or awkward for frequent use, and preferable for more sporadic interaction, where it can provide faster or more convenient access to instant control than other devices or modalities.

Gaze+Hold depends on closing of one eye while using the other as a behaviour that does not occur naturally and that there can be used as a robust signal of intent. This is beneficial in clearly delimiting gaze input actions without need for any other modality but it also presents significant limitations, as both vision and gaze control are reduced to one eye for the duration of the interaction. Part of the visual periphery become unavailable, and gaze input is

slower, not as accurate and more demanding compared to binocular gaze. Users are nonetheless effective with Gaze+Hold, even for a highly repetitive task as used in our study. Perceived effort may be lower when the task is less intensive. However, more complex input composed of multiple steps (e.g. multi-selection, or menu navigation) will require prolonged closure of one eye, and may be experienced as more fatiguing – our study did not assess this. It is also known that not everybody is able to wink with both eyes, although in our study we found all users able to use the technique.

From a technical perspective, the eye tracker's accuracy and precision also affect the usability. Our results are based on the use of a consumer-grade remote binocular eye tracker for desktop usage. The techniques presented are equally applicable in other environments where eye tracking is readily available, such as mixed reality head mounted displays, however further study is necessary due to differences in the characteristics of eye trackers. The techniques presented in this paper invite future study of the usability of Gaze+Hold to support interaction in specific applications, or as a part of a holistic eyes-only interface.

## 7 CONCLUSION

Gaze+Hold enables eyes-only direct manipulation by using closure of one eye to modulate gaze input performed with the other. The technique is unique in leveraging the eyes as separate channels for complementary input. One-eyed gaze is not as fast as binocular gaze, but closure of one eye is highly effective for controlling mode of interaction and delimiting gaze input, without need for any other device or modality.

Our work shows that users are effective with Gaze+Hold, and that they perform equally well with dominant and non-dominant eye. This is significant as it allows both eyes to be utilised flexibly, for access to different input modes as well as dynamic configuration of eyes-only control. Gaze+Hold thus opens up a new design space for eyes-only techniques and interfaces, for which this work provides a foundation and key design considerations.

## REFERENCES

- Mihai Băce, Teemu Leppänen, David Gil de Gomez, and Argenis Ramirez Gomez. 2016. UbiGaze: Ubiquitous Augmented Reality Messaging Using Gaze Gestures. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications* (Macau) (SA '16). Association for Computing Machinery, New York, NY, USA, Article 11, 5 pages. <https://doi.org/10.1145/2999508.2999530>
- H Collewijn, J Van Der Steen, and RM Steinman. 1985. Human eye movements associated with blinks and prolonged eyelid closure. *Journal of neurophysiology* 54, 1 (1985), 11–27.
- Stanley Coren. 1999. Sensorimotor Performance as a Function of Eye Dominance and Handedness. *Perceptual and Motor Skills* 88, 2 (1999), 424–426. <https://doi.org/10.2466/pms.1999.88.2.424>
- Alexander De Luca, Roman Weiss, and Heiko Drewes. 2007. Evaluation of eye-gaze interaction methods for security enhanced PIN-entry. In *Proceedings of the 19th australasian conference on computer-human interaction: Entertaining user interfaces*. 199–202.
- Heiko Drewes and Albrecht Schmidt. 2007. Interacting with the computer using gaze gestures. In *IFIP Conference on Human-Computer Interaction*. Springer, 475–488.
- Tomer Elbaum, Michael Wagner, and Assaf Botzer. 2017. Cyclopean vs. dominant eye in gaze-interface-tracking. (2017).
- Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 457–466.
- Pedro Figueiredo and Manuel J. Fonseca. 2018. EyeLinks: A Gaze-Only Click Alternative for Heterogeneous Clickables. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (Boulder, CO, USA) (ICMI '18). ACM, New York, NY, USA, 307–314. <https://doi.org/10.1145/3242969.3243021>

- Rainhard Dieter Findling, Tahmid Quddus, and Stephan Sigg. 2019. Hide my Gaze with EOG! Towards Closed-Eye Gaze Gesture Passwords that Resist Observation-Attacks with Electrooculography in Smart Glasses. In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia*. 107–116.
- Agostino Gibaldi, Mauricio Vanegas, Peter J Bex, and Guido Maiello. 2017. Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behavior research methods* 49, 3 (2017), 923–946.
- Argenis Ramirez Gomez and Hans Gellersen. 2018. Smooth-i: smart re-calibration using smooth pursuit eye movements. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ACM, 10.
- Kristen Grauman, Margrit Betke, Jonathan Lombardi, James Gips, and Gary R Bradski. 2003. Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. *Universal Access in the Information Society* 2, 4 (2003), 359–373.
- Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
- Henna Heikkilä and Kari-Jouko Riihinen. 2012. Simple gaze gestures and the closure of the eyes as an interaction technique. In *Proceedings of the symposium on eye tracking research and applications*. 147–154.
- Fabian Hemmert, Danijela Djokic, and Reto Wettach. 2008. Perspective Change: A System for Switching between on-Screen Views by Closing One Eye. In *Proceedings of the Working Conference on Advanced Visual Interfaces (Napoli, Italy) (AVI '08)*. Association for Computing Machinery, New York, NY, USA, 484–485. <https://doi.org/10.1145/1385569.1385668>
- Jutta Hild, Patrick Petersen, and Jürgen Beyerer. 2016. Moving Target Acquisition by Gaze Pointing and Button Press Using Hand or Foot. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications (Charleston, South Carolina) (ETRA '16)*. ACM, New York, NY, USA, 257–260. <https://doi.org/10.1145/2857491.2857535>
- T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey. 1989. Human-computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics* 19, 6 (1989), 1527–1534.
- Howell Istance, Richard Bates, Aulikki Hyrskykari, and Stephen Vickers. 2008. Snap Clutch, a Moded Approach to Solving the Midas Touch Problem. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications (Savannah, Georgia) (ETRA '08)*. Association for Computing Machinery, New York, NY, USA, 221–228. <https://doi.org/10.1145/1344471.1344523>
- Robert J. K. Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people (Seattle, Washington, United States) (CHI '90)*. ACM, New York, NY, USA, 11–18. <https://doi.org/10.1145/97243.97246>
- Ricardo Jota and Daniel Wigdor. 2015. Palpebrae Superioris: Exploring the Design Space of Eyelid Gestures. In *Proceedings of the 41st Graphics Interface Conference (Halifax, Nova Scotia, Canada) (GI '15)*. Canadian Information Processing Society, CAN, 273–280.
- Arie E Kaufman, Amit Bandopadhyay, and Bernard D Shaviv. 1993. An eye tracking computer user interface. In *Proceedings of 1993 IEEE Research Properties in Virtual Reality Symposium*. IEEE, 120–121.
- Mohamed Khamis, Mariam Hassib, Emanuel von Zezschwitz, Andreas Bulling, and Florian Alt. 2017. GazeTouchPIN: protecting sensitive data on mobile devices using secure multimodal authentication. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. 446–450.
- Piotr Kowalczyk and Dariusz Sawicki. 2019. Blink and Wink Detection as a Control Tool in Multimodal Interaction. *Multimedia Tools Appl.* 78, 10 (May 2019), 13749–13765. <https://doi.org/10.1007/s11042-018-6554-8>
- Pin-Sung Ku, Te-Yen Wu, Ericka Andrea Valladares Bastias, and Mike Y. Chen. 2018. Wink It: Investigating Wink-Based Interactions for Smartphones. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (Barcelona, Spain) (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 146–150. <https://doi.org/10.1145/3236112.3236133>
- Manu Kumar, Terry Winograd, and Andreas Paepcke. 2007. Gaze-Enhanced Scrolling Techniques. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems (San Jose, CA, USA) (CHI EA '07)*. Association for Computing Machinery, New York, NY, USA, 2531–2536. <https://doi.org/10.1145/1240866.1241036>
- Michael Land and Benjamin Tatler. 2009. *Looking and acting: vision and eye movements in natural behaviour*. Oxford University Press.
- Chris Lankford. 2000. Effective Eye-Gaze Input into Windows. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications (Palm Beach Gardens, Florida, USA) (ETRA '00)*. Association for Computing Machinery, New York, NY, USA, 23–27. <https://doi.org/10.1145/355017.355021>
- Christof Lutteroth, Moiz Penkar, and Gerald Weber. 2015. Gaze vs. mouse: A fast and accurate gaze-only click alternative. In *Proceedings of the 28th annual ACM symposium on user interface software & technology*. 385–394.
- Diako Mardanbegi, Tobias Langlotz, and Hans Gellersen. 2019. Resolving Target Ambiguity in 3D Gaze Interaction through VOR Depth Estimation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 612.
- JJ Marotta, TS Perrot, D Nicolle, P Servos, and MA Goodale. 1995. Adapting to monocular vision: grasping with one eye. *Experimental Brain Research* 104, 1 (1995), 107–114.
- Eric Missimer and Margrit Betke. 2010. Blink and wink detection for mouse pointer control. In *Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments*. ACM, 23.
- Marcus Nyström, Richard Andersson, Kenneth Holmqvist, and Joost Van De Weijer. 2013. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior research methods* 45, 1 (2013), 272–288.
- Lyelle L Palmer. 1976. Inability to wink an eye and eye dominance. *Perceptual and motor skills* 42, 3 (1976), 825–826.
- Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: combining gaze with multi-touch for interaction on the same surface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 509–518.
- Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze+pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*. 99–108.
- Marco Porta, Alice Ravarelli, and Giovanni Spagnoli. 2010. ceCursor, a contextual eye cursor for general pointing in windows environments. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. 331–337.
- Argenis Ramirez Gomez and Hans Gellersen. 2019a. Exploring the Sensed and Unexpected: Not Looking in Gaze Interaction. In *Proceedings of the Halfway to the Future Symposium 2019 (Nottingham, United Kingdom) (HTTF 2019)*. Association for Computing Machinery, New York, NY, USA, Article 35, 7 pages. <https://doi.org/10.1145/3363384.3363479>
- Argenis Ramirez Gomez and Hans Gellersen. 2019b. SuperVision: Playing with Gaze Aversion and Peripheral Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 473.
- Argenis Ramirez Gomez and Hans Gellersen. 2020. More than Looking: Using Eye Movements Behind the Eyelids as a New Game Mechanic. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 362–373.
- Javier San Agustín, Julio C Mateo, John Paulin Hansen, and Arantxa Villanueva. 2009. Evaluation of the potential of gaze input for game interaction. *PsychNology Journal* 7, 2 (2009), 213–236.
- Robin Shaw, Everett Crisman, Anne Loomis, and Zofia Laszewski. 1990. The eye wink control interface: using the computer to provide the severely disabled with increased flexibility and comfort. In *[1990] Proceedings. Third Annual IEEE Symposium on Computer-Based Medical Systems*. IEEE, 105–111.
- Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, and Hans Gellersen. 2020. Outline Pursuits: Gaze-Assisted Selection of Occluded Objects in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376438>
- Ludwig Sidenmark and Hans Gellersen. 2019. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 27, 1, Article 4 (Dec. 2019), 40 pages. <https://doi.org/10.1145/3361218>
- HJ Simonsz and LJ Bour. 1991. Covering one eye in fixation-disparity measurement causes slight movement of fellow eye. *Documenta Ophthalmologica* 78, 3-4 (1991), 141–152.
- Hari Singh and Jaswinder Singh. 2018. Real-time eye blink and wink detection for object selection in HCI systems. *Journal on Multimodal User Interfaces* 12, 1 (2018), 55–65.
- Oleg Špakov and Dariusz Miniutas. 2005. Gaze-based selection of standard-size menu items. In *Proceedings of the 7th international conference on Multimodal interfaces*. ACM, 124–128.
- Sophie Stellmach and Raimund Dachsel. 2013. Still looking: investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. In *Proceedings of the sigchi conference on human factors in computing systems*. ACM, 285–294.
- Jayson Turner, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. Gaze+RST: integrating gaze and multitouch for remote rotate-scale-translate tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 4179–4188.
- Jayson Turner, Eduardo Velloso, Hans Gellersen, and Veronica Sundstedt. 2014. EyePlay: applications for gaze in games. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*. ACM, 465–468.
- Eduardo Velloso, Marcus Carter, Joshua Newn, Augusto Esteves, Christopher Clarke, and Hans Gellersen. 2017. Motion correlation: Selecting objects by matching their movement. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 3 (2017), 1–35.
- Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct control of ambient devices by gaze. In *Proceedings of the 2016 acm conference on designing interactive systems*. 812–817.
- Roel Vertegaal. 2008. A Fitts Law comparison of eye tracking and manual input in the selection of visual targets. In *Proceedings of the 10th international conference on Multimodal interfaces*. ACM, 241–248.

- Mérodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Zurich, Switzerland) (UbiComp '13)*. ACM, New York, NY, USA, 439–448. <https://doi.org/10.1145/2493432.2493477>
- All About Vision. 2017. Finding Your Dominant Eye Video. <https://www.allaboutvision.com/video/dominant-eye.htm>
- Bryan Wang and Tovi Grossman. 2020. BlyncSync: Enabling Multimodal Smart-watch Gestures with Synchronous Touch and Blink. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376132>
- Colin Ware and Harutune H Mikaelian. 1986. 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*. 183–188.
- Jacob O Wobbrock, James Rubinstein, Michael W Sawyer, and Andrew T Duchowski. 2008. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. 11–18.
- Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (Denver, Colorado, USA) (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2878–2889. <https://doi.org/10.1145/3025453.3025790>
- Xuan Zhang and I. Scott MacKenzie. 2007. Evaluating Eye Tracking with ISO 9241 - Part 9. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 779–788.
- Yanxia Zhang, Jörg Müller, Ming Ki Chong, Andreas Bulling, and Hans Gellersen. 2014. GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Seattle, Washington) (UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 559–563. <https://doi.org/10.1145/2632048.2636071>