# Where to display? How Interface Position Affects Comfort and Task Switching Time on Glanceable Interfaces

Samat Imamov*
Center for Human-Computer Interaction
Virginia Tech

Daniel Monzel†
Center for Human-Computer Interaction
Virginia Tech

Wallace S. Lages‡
Center for Human-Computer Interaction
Virginia Tech

Figure 1: The task consists on selecting two highlighted symbols. The first appears on the TV in the beginning of the task (Left). To obtain the second, the participant looks for the corresponding symbol in the glanceable interface (Middle). Back on the TV, the participant press the buttons corresponding to the first and second symbols. Checkbox appear at fixed positions to indicate if the first and second were correct.

## Abstract

A critical decision when designing glanceable information displays is where to place the content. Since blocking the center of the field of view with virtual information is not desirable, designers often opt for placement in the visual periphery. Another option is to only show virtual content when needed. However, no study has been made to systematically evaluate world-locked content position, considering both cognitive and physiological constraints. With this goal in mind, we designed a scenario that mimics context switching between a real world-task and an information display. We then conducted a within-subjects study to evaluate the effect of position, parameterized by horizontal angle, vertical angle, and distance from the user. Our results show that context switching time increases as the information is displayed far from the task position. The same happens with discomfort: content placed at eye level, or below, was faster and more comfortable than in other positions. We also found participants preferred content at medium distances, although they were also faster with content at far distances.

**Index Terms:** Human-centered computing—Mixed and Augmented Reality Human-centered computing—Information Interfaces and Presentation—Miscellaneous

## 1 Introduction

Many dual task scenarios involve the use of a contextual display, i.e., showing information that is spatially and temporally relevant to the task at hand. For example, displaying information about an assembly procedure as the user perform it [12, 13], or directions to the destination while driving a car [11]. In other cases, the information may not be semantically linked to a physical feature in the world, but is still relevant to the task being executed. Checking the current date when talking to someone, or how many minutes are left until your pizza is ready belong to the second category. These scenarios are powerful because they allow us to access digital information while minimizing the overhead of context switching.

When there is a clear connection between the virtual content and the physical world, the virtual information can be displayed directly over the physical object or connected to it by arrows or lines [3]. In this work, we look at the alternate problem: if a small bit of information is necessary, and the display is free, where should it be rendered? As technology progress towards all-day wearable displays, we argue that this scenario will become more common, as users seek to avoid the unnecessary occlusion and distraction caused by copious amounts of information displayed over the field of regard.

Although long and complex information may need to be visible during longer periods of time, in many cases information can be acquired at a glance [4]. Glanceable interfaces are low cognitive effort displays that enable users to 1) maintain focus on the primary task, 2) determine the most appropriate time to switch tasks, and 3) easily reacquire the original task [22]. A few examples would be: checking the current temperature, the tool needed for the next assembly task, or the day of the month. Even in those simple tasks, alternating between an ongoing real-world task and the AR content may involve several steps. Depending of where the content is rendered, it may be necessary to rotate the head and eyes, adjust eyes vergence and accommodation, and shift the focus of attention. Unfortunately, no systematic study have looked into the effect of interface position in task switching time and comfort on glanceable interfaces.

The study described in this paper investigates the effect of horizontal angle, vertical angle, and distance from the user in the completion time of a task that spans between two interfaces. The first interface emulates a primary task, while the second emulates a glanceable AR interface that can be summoned on demand. We compared the time spent away from the primary task for 18 different positions and also solicited participants' input on the positions they found more comfortable using. We found that eye level content was faster and more comfortable than other positions. In overall, both time and discomfort increased with distance from the central position. However, while most participants judged that middle distances were more comfortable, we found that they were also faster with screens further away and closer to the spatial position of the primary task. These and other findings are discussed in more detail later. Next we

*e-mail: samat@vt.edu
†e-mail: drmonzel@vt.edu
‡e-mail:wlages@vt.edu

2642-5254/20/$31.00 ©2020 IEEE
DOI 10.1109/VR46266.2020.00110

present background and related work, followed by a description of our study design. Finally we present our findings and conclusions.

## 2 Background and Related Work

### 2.1 Context Switching

Attending to different points in space present several challenges. First, the eyes need to rotate so that the brain can able fuse the images received by each eye. In addition, the eyes must adjust their focal lengths so that the image on the retina becomes focused. Also, pupil diameter may need to be adjusted, since it affects the depth of field [37]. Normally, those three are linked together in what is known as accommodation-convergence reflex. When the target is not along the principal vision axis, the eyes and the head may need to rotate so that image of the interest point can be projected on the fovea, the area with higher visual acuity. For static objects, the fixation point can be quickly changed by means of saccadic eye movements [9].

Most mixed reality displays though, have a single focal plane that may, or may not coincide with the distance of the virtual object. Fortunately, humans can learn how to disconnect vergence and accommodation process. However even with the a a fixed focal plane all the remaining mechanisms still need to take place.

Our work is also related to research in notification interfaces. During a notification and when accessing a glanceable interface, attention needs to be allocated and directed to different points in space [23]. For example, Rzayev et al. found that notifications that highest number of notifications are missed when participants are performing a concentration-demanding task [34]. Spatial attention seems to be tightly coupled with eye movements: saccadic eye-movements are preceded by a compulsory shift of attention to the saccade goal. Although this happens automatically, it is also possible to pay attention to a location independently of the gaze direction.

However, while notifications typically consist in external interruptions [24], context switching on glanceable interfaces are discretionary [22]. For this reason, the goal of most notification research is to determine when, where, and how to interrupt the user in order no minimize distractions. In most cases, the background task (for example driving) is still running when the interruption occurs (phone message) [17]. Here, we focus on tasks that are completely interrupted when the switch occurs and the glanceable interface is used to continue the primary task.

### 2.2 The Layout of Virtual Information

Prior research has compared information layouts in augmented and virtual reality. Most of them distinguish layouts according to reference frame they are attached to, e.g., user's head, body, or the world [5, 6, 10, 26]. Research in VR has looked into designs that incorporate 2D information with a 3D environments or objects to improve information retrieval [7, 8, 31, 32] or to investigate how relationships between objects and information can enable problem data comprehension and problem solving [27, 28]. Layouts have also been investigated in terms of view management [1, 3, 16], where the goal is optimize the location of multiple AR annotations in the field of view to avoid occlusion and increase understanding.

Not many studies have systematically investigated the effect of information position on task switching time. Rzayev et al. compared text presentation in the top-right, center, and bottom-center position for reading text [35]. Participants read text shown in Rapid Serial Visual Presentation (RSVP) and line-by-line scrolling, while walking and sitting. The authors found that text displayed in the top-right position increased subjective workload and reduced comprehension. Participants preferred the bottom-center and center positions, although they mentioned a potential problem with occlusion. Again using RSVP, Ku et al. explored how text case and font affect the ability to read text at two eccentricity values. The authors found that even though reading was possible at both positions, mental load at 8°was significantly higher than when the text was showed closer to the central part of the retina (5°) [19].

Klose et al. studied how reading in either head-locked or body-locked reference frames, affected three primary tasks (a visual stimulus response task, simple walking, walking task with obstacles) [18]. Two heights were tested: 10 degrees above and 35 degrees below the horizontal line of sight. Participants preferred the head-locked presentation during the visual stimulus test, which did not require them to move. However, in the body-locked condition, the text was presented outside the device field of view, which was clearly disadvantageous. Schankin et al. measured the reaction time to simple color stimuli presented either in AR or in the real world while participants focused their attention either on AR or the world. Participants responded faster to AR and real-world signals when AR signals were presented world-fixed [36].

In summary, previous research has looked into general ways for arranging information or compared a few unique positions, reasonably close to each other. In particular, we could not find any prior work looking into the effect of distance from the user (depth).

### 2.3 Research Questions and Hypothesis

The goal of this research was to understand the impact of position on efficiency and comfort of glanceable interfaces. We sought to answer the following questions: 1-Given a world stabilized glanceable content, where is the most efficient placement for it? 2-How does performance change with depth, vertical, horizontal angles? 3-How does users perceive comfort of using different positions?

Based on previous research, we hypothesized that the time required to switch between two tasks would increase with the distance between them. Shorter distances would reduce the time necessary to change fixation between the targets and would also facilitate the shift in attention between them. Thus, we expected that users would prefer the 3D location in space closest to the position of the primary task.

## 3 Experiment

We designed a study to examine the effect of the position of world-locked interfaces relative to the physical object of primary focus. We chose a task that could only be completed by accessing information displayed in a virtual screen. In this way it was possible to ensure that participants were really using the information from the glanceable display. Our methods consisted in measuring the time spent away from the main task, task errors, and conducting interviews regarding preference and comfort.

### 3.1 Apparatus

For this experiment, participants used a HTC Vive Pro [1] along with an Xbox One wireless controller [2]. The HTC Vive Pro is a wired Virtual Reality (VR) Head-Mounted Display with a resolution of 1440 x 1600 per eye, 90 Hz refresh rate, and a field of view (FOV) of 110 degrees. For this study, we chose using VR instead of AR because current AR technology has a limited FOV. For comparison, Microsoft HoloLens' horizontal field of view is only 30 degrees. Having larger FOV allowed us to place more content inside the screen and better simulate what a future AR headset might provide. This approach is known as Mixed Reality Simulation [20, 29, 33].

The experimental environment was developed using the game engine Unity3D [3]. It consisted of a living room, modeled using Autodesk Maya software [4].

[1] https://www.vive.com/us/product/vive-pro/

[2] https://www.xbox.com/en-US/xbox-one/accessories/controllers/

[3] http://www.unity3d.com

[4] http://www.autodesk.com

### 3.2 Task

Although previous studies have employed reading or visual search as the primary task, we designed a task with lower enough cognitive effort so that it would be equivalent to glanceable interface while still allowing user input. The task consisted of finding a symbol that succeeds a highlighted one by looking into an association table. Each trial consisted of four steps: 1 - four figures were displayed in a TV inside the virtual environment. One symbol was highlighted. The participant memorized the highlighted symbol and pressed the left bumper on the controller to call the glanceable interface. The TV image was turned off at this point. 2- Using the highlighted symbol as a key, the participant looked for its pair on the glanceable information display. Once found, the participant pressed the left bumper on the controller again to dismiss the glanceable display. The previous TV content reappear, without the highlight. 3-The participant press the two buttons corresponding to the direction of the initially highlighted figure and its pair. Green or red check marks appear as a feedback to the participant, ending the trial (Figure 2).

### 3.3 Experimental Design

We used a within-subjects design with three independent variables: vertical angle, horizontal angle, and distance. Since we did not expect left and right positions to yield different results, we arranged the study positions to split the horizontal angle in a non-overlapping way between left and right.

Vertical distance had three levels (0.5 meters above, center, and 0.5 meters below), horizontal position had two levels (0.5 meters, center), and distance had three levels (1m, 2m. 3m). This resulted in (3x2x3) 18 distinct positions. The distances were chosen so that all positions would fit in the HMD field of view, while keeping the content clear and inside a room with reasonable dimensions (5mx5m). The order of presentation was counterbalanced using the combination of the 6 possible positions on the fronto-parallel plane (top-left, top-middle, middle-middle, middle-right, bottom-middle, bottom-right) and the 3 possible depths for a total of 18 different presentation sequences. The primary task was shown at a TV placed 4 meters away from the participant position.

Participants did 12 consecutive trials for each position. The symbol arrangements and correct keys for each trial were pre-defined and rotated between participants using a latin square. The total number of tasks performed by each participant was 216 (18 positions x 12 trials).

To avoid confounding effects caused by perspective (which would make symbols smaller with distance) with the effects we were interested (attention orientation, head movement, eye movement, and vergence) we adjusted the screens so that they would maintain their apparent size from the participant perspective. We believe this decision also has the benefit of making the study more ecologically valid: in a interface designed to be used at distances, symbols and typography would be adjusted so that they would be legible. The background of the glanceable interface was also rendered with an opacity of 78.4% to simulate an AR display (Figure 2-2).

To make sure we measured the diverted from the main task, the study was designed so that the glanceable interface and the task on the TV could not be seen at the same time. Pressing the button to open the interface would automatically hide the task on the TV. Once the interface was dismissed, that task would re-appear on the TV. Thus, the time spent on the glanceable part of the task consisted of two intervals:

*glanceable_time* - time between the first button press summoning of the display and the second causing its dismissal.

*return_time* - the time between the dismissal of the display and the first of two directional button presses.

Here we refer to the sum of both intervals as switching time. It starts with the moment the user is ready to access the information and finishes when he/she is ready to use the information in the task after returning to it. We also gave a detailed opinion-based questionnaire to find out the preferences of participants.

### 3.4 Participants

We recruited twenty-one students from a university (14 male and 7 female) as well as one female faculty member to participate in the study. The mean age was 22.8 years, with a standard deviation of 6.27. Without the faculty member, the mean age was 21.6 years, with a standard deviation of 2.65. Some participants were compensated with course credit for their participation. Participants were required to be English speakers, to be at least 18 years old, and to have normal or corrected vision. These requirements were included in the call for participation emails. Of the 22 participants, 12 participants used glasses or contacts while 10 participants did not need any vision correction. All 22 participants passed a random-dot stereogram (RDS) test given to them before the experiment. The study was approved by the authors university Institutional Review Board. All participation was voluntary and standard IRB precautions were taken to ensure participants' informed consent, anonymity and security during the experiment.

### 3.5 Procedure

Each participant filled out a demographic questionnaire and their interpupillary distance (IPD) was measured using a Sunwin digital pupilometer and adjusted to the best fit in the VR headset. Each participant was administered a random-dot stereotest to assess basic stereoacuity. The experimental procedures were then explained.

Participants sat in a chair wearing a HTC Vive Pro headset and holding an Xbox One controller. Within the virtual world, participants viewed a living room with several 3D objects filling up the space realistically. Within this virtual environment, the series of tasks was administered as discussed in Section 3.2 and Section 3.3. Practice trials were performed to allow participants to gain confidence with using the controller, as well as switching between the virtual television screen and the floating interface. Participants were required a minimum of 3 practice trials but had no maximum limit. The participants were instructed to notify the examiners when they felt comfortable with the task, and most did so after 5 practice trials. After notifying the examiners, the real trials began. Before attempting a task at each interface position, participants were made aware of the location of the interface. Participants were asked to complete 12 trials for each position and a break of 2 minutes was given after every 6 positions, for a total of two breaks. Upon completion of the experiment, a post-experiment questionnaire was conducted while the participants stayed in VR. Each session lasted roughly 45 minutes (10 minutes introduction and briefing, 30 minutes testing, 5 minutes post questionnaire).

## 4 Results

We first present the raw data distribution and then the related statistical analysis. For better statistical power, we considered all individual trial attempts instead of the average of the time for each interface, which would be vulnerable to outliers.

### 4.1 Time

For correct trials, the median time spent switching tasks was 1.44s. The minimum and maximum times observed were 0.61s and 7.9s respectively. The corresponding histogram shows the long right tail, characteristic of time measurements (Figure 4). To avoid including outliers due to distraction and improve the robustness of estimations, we trimmed the data to 5s. This removed approximately 5.8% of the data points.

Figure 5 shows the mean times along the fronto-parallel plane. Participants were slower when using the interface on top left position (A positions on Figure 3, followed by the bottom right position (E

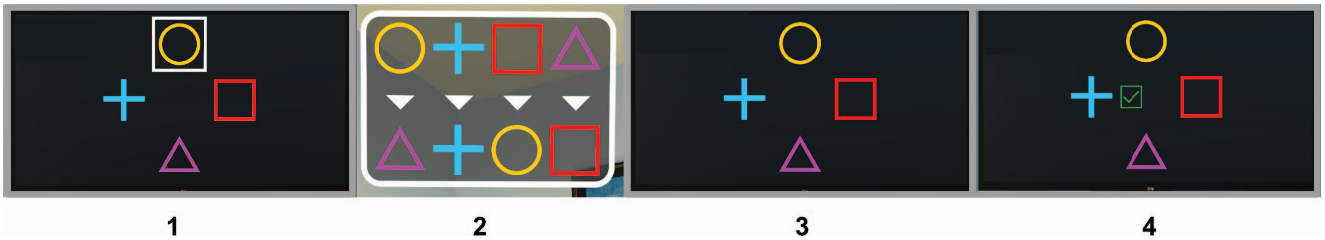


Figure 2: Experimental Task: 1-One symbol is highlighted on the TV; 2- Participant summons the glanceable interface and looks for the pair; 3-Participant dimiss the interface and returns to the TV; 4-Participant press the buttons corresponding to the symbols; Checkbox is not associated with the symbol position.


Figure 3: Arrangement of all positions used in the study with the IDs as used during the post-study interview.

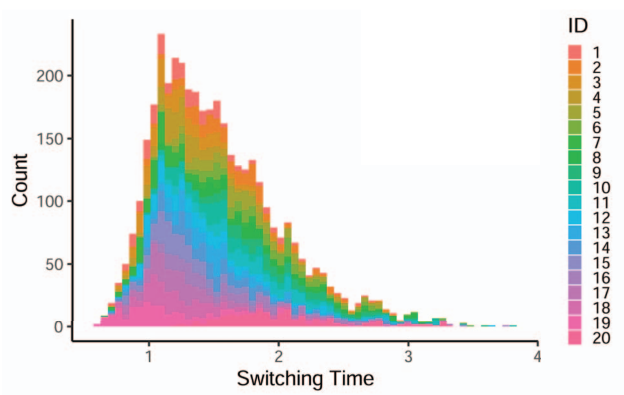


Figure 4: Histogram of Switching Time over all positions. Colors indicate different participants (ID).

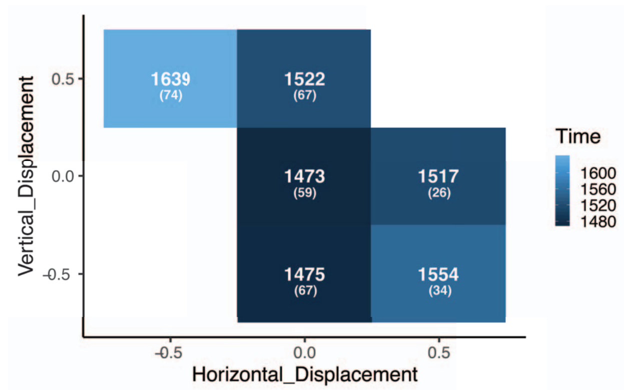


Figure 5: Mean times and standard deviations (in parenthesis) for different vertical and horizontal angles (fronto-parallel plane). Center and lower positions were faster than far up and down. Time milliseconds and distances in meters.

positions on Figure 3). Both central and central-low positions yield faster response.

The mean times across the vertical axis (side projection) are displayed on Figure 6 . Visually, times seem to improve as the interface used was closer to the task position (and away from participants) and lower in the space. The fastest position was the closer to lower end of the TV, while slowest access time were observed on the highest position closer to the user.

Figure 7 shows the mean times across the horizontal axis (top projection). The central positions away from the user (2 and 3 meters) were faster than the remaining positions.

#### 4.1.1 Model Fit

To understand the contribution of distance and angles in the participant's performance we fitted a generalized multiple regression model using the package lme4 in R [2]. For the model, we considered only the absolute value of the horizontal displacement, making the horizontal angle a fully crossed 2-level predictor. We also calculated the relative angles for each interface position. Radians were used to facilitate model identification.

Model selection started by first considering a full model, including the main 3 predictors as fixed effects (horizontal angle, vertical angle, and distance) along the 2-way and 3-way interactions. We used an inverse gaussian model with a log link function to model the time data [21]. The model was then simplified by removing non-significant interactions and comparing the resulting models using Akaiake Information Criterion. The inclusion of interactions did not improve the model, thus we favored the model without interactions to avoid overfitting. The addition of a random term for the participant also resulted a small improvement but was included to better reflect the dependent data from our experimental design. Variance Inflation Factors for the model were small: 1.046 for horizontal angle, 1.000 for vertical angle, and 1.046 for distance. The deviance of the fitted model was 0.186 compared to 0.262 of the null model. The estimates for the fixed effects are listed on Table 1.

The positive coefficients for vertical and horizontal angles indicate that participants spent more time consulting the interface as the distance from the central position increases. The value for horizontal angle was higher, indicating that the impact per unity of angle was higher in the horizontal than on the vertical axis. The distance estimate was the smaller of the three, indicating a small improvement in

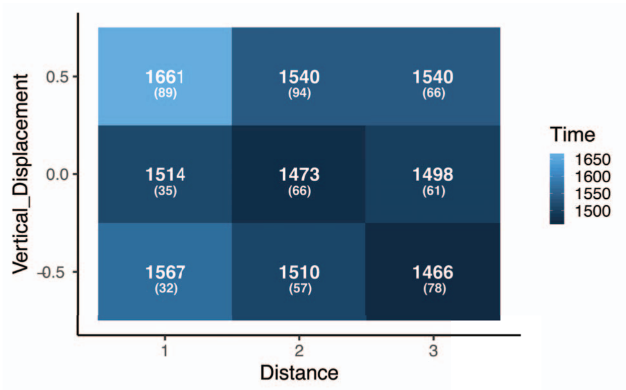


Figure 6: Mean times and standard deviations (in parenthesis) for different elevations and distances. The lower position closer to the TV was faster. The higher position closer to the user was the slowest. Time milliseconds and distances in meters.

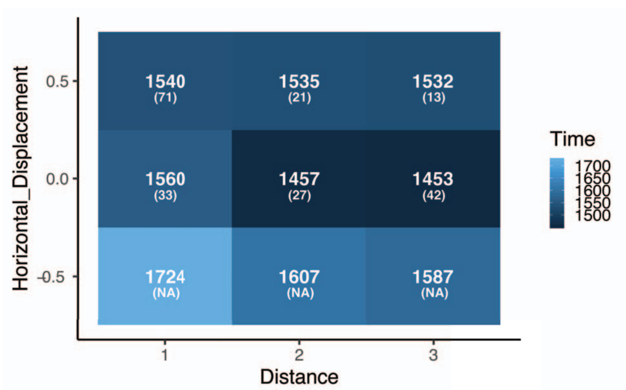


Figure 7: Mean times and standard deviations (in parenthesis) for different horizontal angles and distances. Positions right in front of the participant and closer to the TV were faster. Time milliseconds and distances in meters. Standard deviation is not available for the left side since it only exists at the top level.

switching time as the interface moved away from the user and closer to the TV. The conditional plot for the corresponding model can be seen in Figures 8 and 9.

### 4.2 Accuracy

Overall, participants were very accurate in the tasks. Each participant did 216 trials (12 repetitions on 18 positions). There were only 59 errors in 4320 total total trials (1.4%). Within 17 participants who had at least one mistake, the mean number of errors was 3.4 and the maximum was 16 (a single participant). A Kruskal-Wallis rank sum test on the groups with errors did not detect any significant difference between them ($\chi^2 = 16$, df=16, p_value=0.453). The distribution of errors across the main independent variables can be seen on Figure 10.

### 4.3 Comfort and Preferences

After the trial concluded, participants gave feedback through a verbal questionnaire. Approximately 59.1% of participants preferred the screens at 2 meters, as opposed to 31.8% preferring a farther distance (3 meters) and 9.1% preferring a closer distance (1 meter). Participants also preferred the screens with smaller viewing angle from the television to avoid turning their head as much. 63.6% of participants preferred the screens to be placed at medium (eye level) height, as opposed to 31.8% preferring a lower height and 4.6% preferring a higher height. In general, participants preferred looking down compared to any other non-zero angle of head direction. Figure 11 shows a weighted average for the choice participants made for the five best positions in terms of comfort. The preference for central positions is clearly visible, along with a preference for lower positions. For the top-center and lower-center positions, the 3 meters distance was also more preferred than the closest one (1 meter).

Table 1: Fixed Effects for model mixed model
$T \sim Horizontal\_Angle + Vertical\_Angle + Distance + (1||ID)$

| | | | | |
|---|---|---|---|---|
| H Angle<br>0.1537 | std.error<br>0.0296 | conf.low<br>0.0956 | conf.high<br>0.2115 | p value<br>2.076e-07 |
| V Angle<br>0.0766 | std.error<br>0.0176 | conf.low<br>0.0420 | conf.high<br>0.1112 | p value<br>1.442e-05 |
| Distance<br>-0.0218 | std.error<br>0.0049 | conf.low<br>-0.0313 | conf.high<br>-0.0122 | p value<br>7.988e-06 |

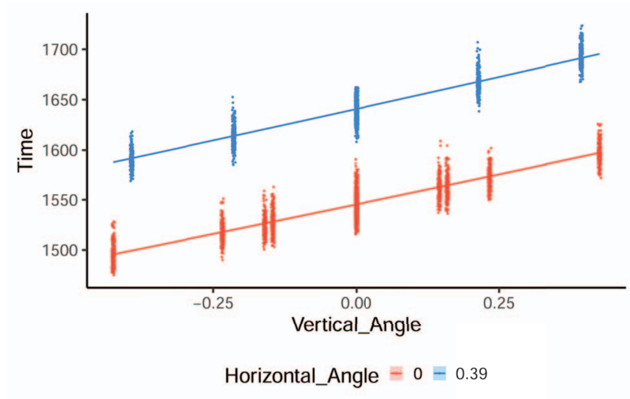


Figure 8: Time vs Vertical Angle. Participants were faster with lower vertical angles and with central view (horizontal angle zero)

Several participants preferred screens that did not conflict with environmental details: *"When it's right on the TV, it's easy to discern the information coming from there. The others don't have much behind them to distract. Some of the lower ones I got distracted by the crease in the carpet"* (P17), *"[there] wasn't too much behind 1d and 2d, similar with 1b 2b and 3b. Although the shadow from the TV was a little distracting but not nearly as much as the other stuff"* (P19), *"the light highlighted the a's [group] better"* (P12), *"The ones I chose I thought were easiest because they were up against solid backgrounds (the bluish wall and the black TV screen) so it took less time for me to distinguish them and process"* (P22).

Participants preferred looking down compared to any other non-zero angle of head direction. One possible reason they gave for this included familiarity with cell phone usage: *"Glancing down feels natural to me, I'm already glancing down at things like my phone"* (P7), *"... I also liked the lower ones because it was similar to checking my phone and I was used to it"* (P15).

Participants stated that they struggled more to focus on the screens at a close distance, while the farther ones were easier to adjust to visually: *"Distance played a big part, I didn't like the ones that were too close to me"* (P11), *"it was harder for me to focus on the ones that were closer since I have a bit of astigmatism"* (P13), *"... the ones farther away were easier to see and grasp right away. The ones that were closest to [the TV] screen without being directly in front of screen were good too"* (P3).

Some participants noted that the medium distance was more

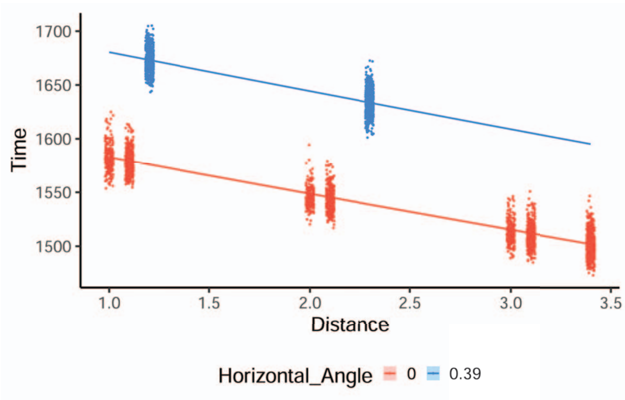


Figure 9: Time vs Distance. Participants were faster with interfaces centered in front of them and at farther distances.

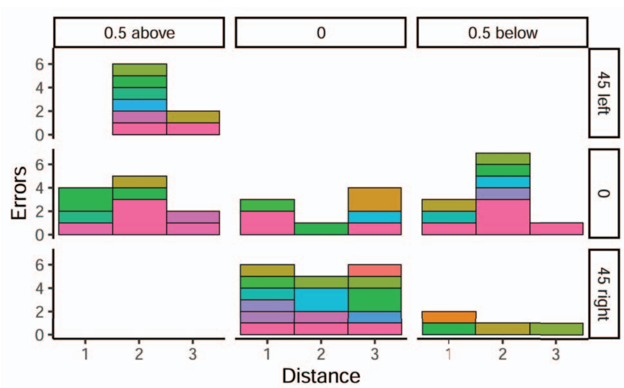


Figure 10: Number of errors for each of the experiment conditions. Only 17 participants made mistakes and most only made one mistake per condition. Colors indicate different participants.

comfortable because it provided a spacial distinction between the floating screen and the television: *"The distance between the screens is good for my eyes, not too close to me or too close to TV, and the distance is comfortable"* (P18), *"The ones in the middle were easy to check but I didn't like the one far away in the middle because it occupied the same visual space as the TV and that confused me"* (P15).

## 5 Discussion

As expected, completion time increased and the comfort rating decreased with vertical and horizontal displacement. However, when looking at task completion time and rankings, we see that positions located at left and right sides had significantly worse rankings and higher completion times than the others. This result is surprising. Prior research has found that the perceived exertion rates for vertical and horizontal gaze angles are practically symmetrical up to 40 degrees (in our study, interfaces were within 24 degrees) [25]. The magnitude of the difference between vertical and horizontal is better seen in the model fit: the coefficient of the horizontal predictor is 0.15 while the vertical coefficient is only 0.08. One explanation for this result is that the HMD weight might have modified the head's center of gravity in a way that affected lateral movements more than vertical ones.

Within the positions with zero horizontal displacement, we observed that those located at or below the center were rated better than the top ones. Again, the comfort rankings seem to agree with the

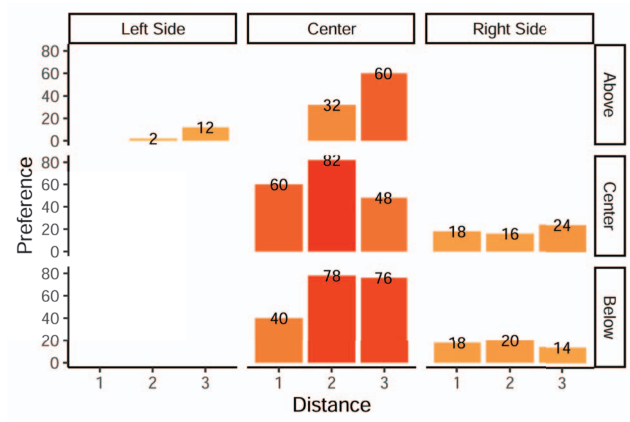


Figure 11: Weighted preferences of position by the participants regarding comfort. Most participants preferred center and below-center positions at distances 2m and 3m.

measured completion times. On the fronto-parallel plane, this result is similar to the findings of studies on regular computer screens, which found that gaze directions from 10 to 25 °downwards are more comfortable at shorter distances [15,25,38]. However, when looking at increasing distances, our result points to a decrease in time rather than increase. Since the lower look seems to favor some vergence and accommodation mechanisms [14,30], it is possible that the result was modified by the decoupling between accommodation and vergence required to use the HMD. Further research is necessary to investigate this point.

When looking at the 3 best positions, the match between weighted preferences and time completion is no longer perfect. The positions judged as more comfortable were, in order: central at 2 meters, and lower-central at 2 and 3 meters. However, participants were faster with the lower-central at 3 meters, followed by the central at 2 meters and central at 3 meters. A Kruskal-Wallis Rank test rejected the hypothesis of no significant difference between the first 3 groups ($\chi^2 = 12.299$, p=0.01526). Thus, the disagreement might be due an imprecision in judgment, ranking weights, or to an external factor that was not captured in this study (e.g., mental load).

## 6 Limitations and Future Work

Our study focused on a narrow use case of glanceable interfaces, in order to investigate a large set of positions. We also avoided variables that could confound our results, which reduces some of the generality of the findings. Future studies could address some of these limitations by allowing participants to walk freely and perform tasks at more than a single point. In this way it could be possible to know if the preferences we observed are relative to the user or relative to the task. It could also help to probe any difference between world stabilized vs head stabilized interfaces. On the other hand, we did not control or record eye gaze. It is possible that some users adopted different strategies on how much to rotate their head or eyes.

Another possible area for future work is to investigate different tasks. Our task was purposefully simple so that it could simulate an interface with glanceable content and still have some interaction that would allow us to ensure that information was indeed assimilated. If that is not necessary, more natural tasks can be designed. Since we have seen a correlation between task performance and comfort, it may be possible to evaluate new tasks by solely looking into comfort metrics. Finally, we investigated distances starting from the user position to before the task focus (TV).

Regarding depth, our study only covered positions between the user position and the task focus (TV). In a future study, it would be valuable to investigate interface positions that beyond the task focus. For example, in a environment similar to the one used, the glanceable content could be positioned on the wall behind the TV. We hypothesize that in a real AR application (where the focal distance changes), placing the content far away would be slightly more comfortable and faster due to relaxation of eye's lenses.

Finally, the background of our glanceable interfaces was semi-transparent, which allowed the user to see some of room through behind them. Although we adjusted the opacity to keep the content clearly visible at all positions, practical AR implementations may suffer from lack of contrast, specially if used outside. In those cases, consideration of the best interface position may need to include background characteristics in addition to the variables investigated in this paper.

## 7 Conclusions

We investigated 18 positions in space where glanceable interfaces can be positioned to display content relative to a point of interest. Using a task that ensured accurate time for spent during switching, we evaluated both time and comfort. We found that participants spent less time to complete the tasks when interfaces were positioned centrally, below the task, and away from the user (thus, closer to the task). The time results partially matched preferences regarding comfort: participants indicated slight more preference for the middle position, instead of the farther one. The study suggests designers should avoid placing interfaces too close or away from the axis between the user and the task.

## Acknowledgments

This work was supported in part by a Google Research Award.

## References

[1] R. Azuma and C. Furmanski. Evaluating label placement for augmented reality view management. *Proceedings - 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2003*, pp. 66–75, 2003. doi: 10.1109/ISMAR.2003.1240689

[2] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01

[3] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. *UIST (User Interface Software and Technology): Proceedings of the ACM Symposium*, pp. 101–110, 2001. doi: 10.1145/502360.502363

[4] B. Bell, S. Feiner, and T. Höllerer. Visualization viewpoints: Information at a glance. *IEEE Computer Graphics and Applications*, 22(4):6–9, 2002. doi: 10.1109/MCG.2002.1016691

[5] M. Billinghurst, J. Bowskill, N. Dyer, and J. Morphett. Evaluation of wearable information spaces. *Proceedings - Virtual Reality Annual International Symposium*, pp. 20–27, 1998. doi: 10.1109/vrais.1998.658418

[6] D. A. Bowman, C. North, J. Chen, N. F. Polys, P. S. Pyla, and U. Yilmaz. Information-rich virtual environments: Theory, tools, and research agenda. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, Part F1290:81–90, 2003. doi: 10.1145/1008653.1008669

[7] A. Cockburn. Revisiting 2D vs 3D Implications on Spatial Memory. *AUIC '04 Proceedings of the 5th Conference on Australasian User Interface - Volume 28*, 28:25–31, 2004.

[8] A. Cockburn and B. McKenzie. Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments. *Conference on Human Factors in Computing Systems - Proceedings*, 4(1):203–210, 2002. doi: 10.1145/503411.503413

[9] A. Eiberger, P. O. Kristensson, S. Mayr, M. Kranz, and J. Grubert. Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display. 2019. doi: 10.1145/3357251.3357588

[10] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the World: 2D windows for 3D augmented. *UIST 1993 - Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, pp. 145–155, 1993. doi: 10.1145/168642.168657

[11] J. L. Gabbard, G. M. Fitch, and H. Kim. Behind the glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE*, 102(2):124–136, 2014. doi: 10.1109/JPROC.2013.2294642

[12] S. J. Henderson and S. Feiner. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. *Science and Technology Proceedings - IEEE 2009 International Symposium on Mixed and Augmented Reality, ISMAR 2009*, pp. 135–144, 2009. doi: 10.1109/ISMAR.2009.5336486

[13] S. J. Henderson and S. K. Feiner. Augmented reality in the psychomotor phase of a procedural task. *2011 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011*, pp. 191–200, 2011. doi: 10.1109/ISMAR.2011.6092386

[14] H. Heuer and D. A. Owens. Vertical gaze direction and the resting posture of the eyes. *Perception*, 18(3):363–377, 1989.

[15] W. Jaschinski, H. Heuer, and H. Kylian. A procedure to determine the individually comfortable position of visual displays relative to the eyes. *Ergonomics*, 42(4):535–49, Apr 1999.

[16] S. Julier, M. Lanzagorta, Y. Baillot, L. Rosenblum, S. Feiner, T. Höllerer, and S. Sestito. Information filtering for mobile augmented reality. *Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2000*, pp. 3–11, 2000. doi: 10.1109/ISAR.2000.880917

[17] S. Kim, J. Chun, and A. K. Dey. Sensors know when to interrupt you in the car: Detecting driver interruptibility through monitoring of peripheral interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 487–496. ACM, 2015.

[18] E. M. Klose, N. A. Mack, J. Hegenberg, and L. Schmidt. Text presentation for augmented reality applications in dual-task situations. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, pp. 636–644, 2019. doi: 10.1109/VR.2019.8797992

[19] P.-S. Ku, Y.-C. Lin, Y.-H. Peng, and M. Y. Chen. Peritext: Utilizing peripheral vision for reading text on augmented reality smart glasses. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 630–635. IEEE, 2019.

[20] B. Laha, D. A. Bowman, and J. D. Schiffbauer. Validation of the mr simulation approach for evaluating the effects of immersion on visual analysis of volume data. *Visualization and Computer Graphics, IEEE Transactions on*, 19(4):529–538, 2013.

[21] S. Lo and S. Andrews. To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6:1171, 2015. doi: 10.3389/fpsyg.2015.01171

[22] T. Matthews. Designing and evaluating glanceable peripheral displays. p. 343, 2006. doi: 10.1145/1142405.1142457

[23] D. S. McCrickard and C. M. Chewar. Designing attention-centric notification systems: five hci challenges. In *Cognitive Systems*, pp. 77–100. Psychology Press, 2006.

[24] D. S. McCrickard, M. Czerwinski, and L. Bartram. Introduction: design and evaluation of notification user interfaces. *International Journal of Human-Computer Studies*, 58(5):509–514, 2003.

[25] M. Menozzi, A. von Buol, H. Krueger, and C. Miège. Direction of gaze and comfort: discovering the relation for the ergonomic optimization of visual tasks. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 14(4):393–9, Oct 1994.

[26] T. Müller and R. Dauenhauer. A Taxonomy for Information Linking in Augmented Reality. In L. T. De Paolis, , and A. Mongelli, eds., *Augmented Reality, Virtual Reality, and Computer Graphics*, vol. 1, pp. 382–401. Springer International Publishing, 2016. doi: 10.1007/978-3-319-40621-3

[27] N. F. Polys, S. Kim, and D. A. Bowman. Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. *Computer Animation and Virtual Worlds*, (November):19–38, 2007. doi: 10.1002/cav

[28] E. D. Ragan, D. A. Bowman, and K. J. Huber. Supporting cognitive processing with spatial information presentations in virtual environments. *Virtual Reality*, 16(4):301–314, 2012. doi: 10.1007/s10055-012-0211-8

[29] D. Ren, T. Goldschwendt, Y. Chang, and T. Höllerer. Evaluating wide-field-of-view augmented reality with mixed reality simulation. In *2016 IEEE Virtual Reality (VR)*, pp. 93–102. IEEE, 2016.

[30] P. H. Ripple. Variation of accommodation in vertical directions of gaze. *American Journal of Ophthalmology*, 35(11):1630–1634, 1952.

[31] G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel, and M. van Dantzich. Data Mountain: Using spatial memory for document management. *UIST (User Interface Software and Technology): Proceedings of the ACM Symposium*, pp. 153–162, 1998.

[32] G. Robertson, M. Van Dantzich, D. Robbins, M. Czerwinski, K. Hinckley, K. Risden, D. Thiel, and V. Gorokhovsky. The task gallery: A 3D window manager. *Conference on Human Factors in Computing Systems - Proceedings*, 2(1):494–501, 2000. doi: 10.1145/332040.332482

[33] M. Rodrigue, A. Waranis, T. Wood, and T. Höllerer. Mixed reality simulation with physical mobile display devices. In *2015 IEEE Virtual Reality (VR)*, pp. 105–110. IEEE, 2015.

[34] R. Rzayev, S. Mayer, C. Krauter, and N. Henze. Notification in vr: The effect of notification placement, task and environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 199–211, 2019.

[35] R. Rzayev, P. W. Woźniak, T. Dingler, and N. Henze. Reading on Smart glasses: The effect of text position, presentation type and walking. *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April, 2018. doi: 10.1145/3173574.3173619

[36] A. Schankin, D. Reichert, M. Berning, and M. Beigl. The Impact of the Frame of Reference on Attention Shifts between Augmented Reality and Real-World Environment. *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*, pp. 25–30, 2017. doi: 10.1109/ISMAR-Adjunct.2017.24

[37] B. Wang and K. J. Ciuffreda. Depth-of-focus of the human eye: Theory and clinical implications. *Survey of Ophthalmology*, 51(1):75–85, 2006. doi: 10.1016/j.survophthal.2005.11.003

[38] E. Woo, P. White, and C. Lai. Ergonomics standards and guidelines for computer workstation design and the impact on users' health–a review. *Ergonomics*, 59(3):464–475, 2016.