Guiding Attention to Out-of-View Targets in Locate and Assembly Tasks Through Visual Cues in Virtual Reality

Daniel Johansen
djohan14@student.aau.dk

Alex Mikkelsen
amikke13@student.aau.dk

Allan Schjørring
aschja14@student.aau.dk

ABSTRACT

Virtual or Augmented Reality glasses is finding its way into the professional life in more and more industries. A problem in locating and assembly tasks is finding out of view objects. Guiding a user's attention to out-of-view targets through visual cues can solve this problem. We suggest that there is no one correct answer, and that the context and task at hand has an impact on which cue is most effective in each situation. The two studies in this paper looks into what visual cues work best in regards to time, cognitive workload and preference in two different contexts with different tasks. Our results suggest *leading* cues have an advantage over *directing* cues in locating and assembly tasks.

INTRODUCTION

In an assembly task, it may be difficult to find the right target for the task if it is out of the user's field of view (FoV) or if it is among other objects in a *visually noisy* environment. A solution to this problem is the use of *visual cues* to guide user attention to a specific target using technology such as AR glasses. The visual cues help the user by either *leading* or *directing* the user's attention to the target. The use of visual cues frees up the user's resources in terms of cognitive load (e.g. in visual search) and performance (e.g. acquisition time). In related work, we found that users did two types of tasks: a 2 Degrees-of-Freedom (DoF) locating task or a 3 DoF assembly task. These two types were different in terms of *depth*. In the 2 DoF locating task, the user located a target in a 360° environment regardless of depth. In the 3 DoF assembly task the user both located and interacted with a target in 3D space. Furthermore, the studies also had different kinds of cues that conveyed direction and/or depth in their own way. The differences meant that a cue either *lead* or *directed* the attention towards the target. There were also differences in how studies conveyed 3D space and direction in their cues - for example by projecting the 3D space in to a 2D plane, or by projecting a 3D path to the target into a 2D cue. We found that related work did not examine cues in both 2 DoF and 3 DoF tasks. Instead they only examined what cues worked best in their specific context, e.g. a locating task [7], or birdhouse [9] assembling task. Furthermore, no studies examined both *leading* and *directing* cues in the same context in 2 DoF locating tasks. Related work suggests that a *leading* cue such as a *Wedge* guides a user to an out-of-view target fastest in terms of Degrees-Per-Second in 2 DoF locating tasks, whereas in 3 DoF tasks both cue types seem to perform the same. This leads us to the following statement:

How do cues guide the users in tasks with differing complexity and which cues are best in those tasks?

Based on that, the goal of this paper is to find the right cue for the right context in terms of user performance, and perceived workload. Related work suggests that *leading* cues score lower in NASA-TLX questionnaire. It is also important to figure out how distracting visual cues can be when the user is performing a task. The goal is not to distract the users from their task, but to improve it.

RELATED WORK

In order to understand the results from related work, we have explained the theory behind *attention* and *visual cues* first. After that, we analysed the results from previous work from studies that use cues in 2 DoF and 3 DoF tasks. In order to know how visual cues have an effect on users, it is useful to know how the human visual system works. This information gives us an understanding in how cues are utilised in terms of placement, movement, etc.

Human Visual System

The human FoV includes around 180°, the green part in Figure 1. The foveal vision consists of 1-2° where the vision is the sharpest [11], the brown cone. This part of the vision is best at detecting colour and shapes. There is also a *Useful FoV* (UFoV) which is the visual angle, between 5-15°, where one is still able to gather information quickly without moving one's eye [11], the blue cone in the figure. Outside the UFoV lies the near peripheral vision, up to 30°. The rest is the far peripheral vision. The peripheral vision works best at detecting movement and light.

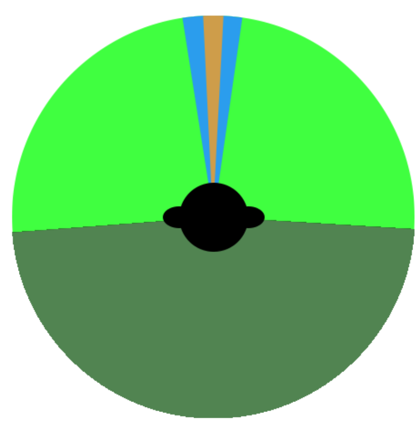


Figure 1: Central vision, brown cone: 1-2°. Useful FoV, blue cone: Up to 15°. Human FoV: Just over 180°. The dark green is out-of-view.

Attention

Human attention includes both perceptual and cognitive processes [3]. The perceptual processes includes what the user sees with the central vision or in the periphery. *Serial* attention is when eye-movement is necessary in order to attend locations in a sequence. *Parallel* attention denotes when a user attends one or more locations without moving their eye, such as driving, crossing a street, or searching for objects. In these cases a person would have their focus on one arbitrary point, but still be attentive to one or more points in the periphery. Generally, parallel attention guides the serial attention.

Endogenous and exogenous attention

Humans have two systems within parallel attention that process and select information. One of the systems is the voluntary part called the *endogenous* attention where a person chooses to monitor a specific location, whereas the other system called the *exogenous* attention is the reflexive part that is non controllable [3].

Visual cues take advantage of these two systems. An example of an exogenous cue is simply a flashing target - a person will by reflex look at it. This is known as pre-attentive processing, which is the theory that explains how the brain processes bottom-up exogenous cues subconsciously [4]. In other words, some elements such as colour, shape, and motion stand out to our brain without us thinking about it. Furthermore, the exogenous cue will only grab the attention temporarily, which takes around 110ms. On the other hand, endogenous cues rely on the knowledge and experience of a person - a top-down cognitive process. An arrow represents an endogenous cue that people learn how to interpret - an arrow points towards an object. Contrary to the exogenous cues, endogenous cues do not need to be exactly on the target, they can be in an arbitrary location, e.g. front of the user's FoV at all times. This is because people can learn how an endogenous cue works. A user can voluntarily attend an endogenous cue for as long as needed, however it takes around 300ms to deploy. A term relevant to endogenous cues is *amodal completion* which is the fact that humans can complete an incomplete shape in their mind [6].

Visual Cues

An example case of a visual cue used to guide the user to a target is in a mobile satellite disk assembly task where a simple *2D arrow* guides the user's attention to the next satellite disk part. In this case the user could wear a Microsoft HoloLens, and the 2D arrow would always be visible in the centre of the FoV.

Leading and directing cues

We define a *leading* cue as being attached to a target, meaning that some part of the cue is always spatially connected to the target. A simple example of this is a line cue that starts from the centre of the target and ends in the periphery of the user's FoV. The user can at all times see the cue and how it responds to head-movements. Another leading cue example is a *Halo* cue [7], which is a circle that has the centre in the target's center. If a target is out-of-view in the right side, the edge of the circle will be visible in the user's right FoV, see Figure 7. If the target is far away from the user's FoV the circle will have a large radius to communicate the distance, which will decrease the closer the user's FoV is to the target.

We define a *directing* cue as a cue that works by giving the user a general direction to the target instead of showing the direct path. The *EyeSee360* cue is an example of a directing cue. The EyeSee360 projects the three dimensional space around a user, into a 2D map that is visible to the user - similar to a real life map. The user's FoV is in the centre of the "map", and the target is a small circular dot, see Figure 6. The user has to learn the spatial relationship between input and output, i.e. how head-movement changes the cue. In this case, if the target is out-of-view on to the left and the user turns to the left, the dot will move right, towards the centre.

Stages of a cue

We have defined four stages of the life time of a cue, which depend on when a target is: *out-of-view*, *entering FoV*, *in central FoV* and *changing target*. Not all cues have different properties in all four stages. When a target is *out-of-view*, a cue visually shows where the target is located. Some cues are situated in the periphery, and some are placed in the centre of view. When the target *enters the FoV*, some cues that have been previously "incomplete" now become complete, e.g. the *Wedge* becomes a triangle, and the *Halo* becomes a circle. When the target is in the *central FoV* some cues disappear completely such as the *Wedge* and *Halo*, or some are still visible such as the dots in the *EyeSee360* and *AroundPlot* [1]. Finally, when a cue changes to the *next out-of-view target*, some will have a noticeable change such as the *Wedge* and *Halo*. These two cues will have a large exogenous effect since they can start by being large, see Figure 7a. The *EyeSee360* and *AroundPlot* will have a smaller exogenous effect since they have small circular dots.

Analysing 2 DoF Tasks

Multiple studies have examined how visual cues perform in 2 DoF VR and MR locating tasks. In these tasks, the users rotated in the pitch and yaw axis in order to locate out-of-view targets [1, 6, 7]. In the three studies, the participant located out-of-view objects that were maximum 180° [1, 6, 7] away from the user's FoV in a simple lab or VR environment with little to no visual noise. Since all out-of-view targets were not directly behind the user, we calculated the mean out-of-view object angle for each study using the MR/VR headset FoV and the equation:

$$\frac{180\ degrees - \frac{FoV}{2}}{2}$$

For example the Oculus Rift had a 110° FoV, which gave the mean out-of-view angle of: 62° - i.e. the out-of-view target was on average 62° outside the user's FoV. Figure 2 shows each platform's FoV and its mean out-of-view angle.

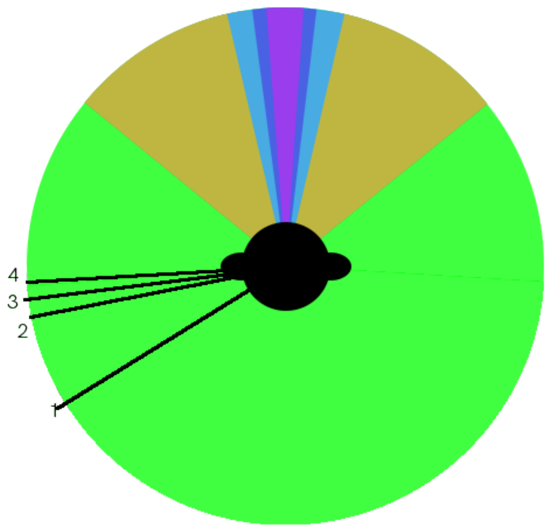


Figure 2: Platform FoVs and Out-Of-View-Angles. *Yellow-Brown*: Oculus Rift 110° (1), *Light Blue*: HoloLens 35° (2), *Dark Blue*: Epson Moverio 23° (3), *Purple*: Google Glass 13° (4). Lines represent the headsets' mean out-of-view angle. *Light-green* is the out-of-view area.

Generally, results showed that in a locating task where the user had to find a specific target, leading cues performed better than directing cues in terms of *Degrees Per Second* (DPS) [1, 6, 7], see Figure 3. DPS is how many degrees the user is turning per second, in the yaw axis. 180 DPS means that the user turns complete around in one second. The DPS was calculated by dividing the mean out-of-view angle by the acquisition times from the studies, e.g. 62°/1.9s = 32°/s. The acquisition time was typically the time it took from the target onset to when the user had the target in the center of their FoV.

Visual noise and type of task

Figure 3 shows that the cues with highest DPS were the *leading* cues *Halo* and *Wedge* with around 32 DPS from the Gruenefeld et al. 2018 study [7]. In the figure, *VR(1)* after *Halo* and *Wedge* means that the cues were used in VR in a locating task were the user had to find one specific target at a time. The figure shows that the more targets there were in the environment (more visual noise), the lower the DPS: *Wedge* and *Halo* cues with *VR(5)* was around 28 DPS and *VR(8)* was around 25 DPS. *Halo*(*VR 1 and 5*) was around 4% faster than *Wedge*(*VR 1 and 5*), while *Halo*(*AR 1 and 5*) was 7-11% faster than *Wedge*(*AR 1 and 5*). This indicates that with a cue that only can move in a small display in AR, a *Halo* cue indicate direction better than a *Wedge* cue. The *Halo*'s circle might convey *amodal completion* better than a *Wedge*. In a bigger VR screen, a *Wedge* seems to convey *amodal completion* more.

The Bork et al. study also support the argument that the number of targets and type of task has an effect on DPS [1]. Bork et al. had three types of locating tasks in MR: *sequential*, *specific*, and *random*. In the *sequential* task, participants located one specific target at a time (the only one visible), whereas in the *specific* task they had to locate one specific target among all other target (eight targets visible). For example, the 3D Radar in the *sequential* task had a DPS of 22 and a DPS of 18 in the *specific* task. On the other hand, many cues with a *random* task had a higher DPS - the 3D Radar had a 32 DPS. In the *random* task, participants had to find any target of the eight - it did not matter which one it was. The results from the Gruenefeld et al. and Bork et al. indicate that the type of task and amount of visual noise has an effect on DPS - more visual noise does not necessarily mean lower DPS but only if the participant had to actively search for a specific target.

Finding a specific target

We compiled a list of all the different cues from the studies in related work that did a 2 DoF locate task and plotted their DPS using our equation from earlier, see Figure 3. The figure shows that cues with the highest DPS were the *leading* cues *Wedge* and *Halo* (used in VR) with the DPS of around 32 [7], whereas the *directing* cues with the highest DPS were EyeSee360 and AroundPlot (used in MR) with around 25 DPS [1]. These results indicate that 1) cues in smaller displays have lower DPS or 2) *leading* cues have higher DPS than *directing* cues. The first statement is supported by the DPS values from the *Halo*(AR) and *Wedge*(AR) cues from the HoloLens tests - they are also around 25 DPS just like the EyeSee360 and AroundPlot DPS values. The EyeSee360 cue gives the participant an overview of the surrounding space - if the cue went from around 35° to around 100° wide, it would take up a lot of visual space and make it harder to get an overview. Since the *Useful FoV* (UFoV) is up to 15°, analysing the visual field with the EyeSee360 and its cues would probably take longer, which could decrease the DPS. The AroundPlot would not take as much visual space as the EyeSee360 since it puts the cues in the periphery and leaves the centre of view clear. The problem could then be that the cues are in the periphery - the cues would perhaps not catch the user's attention, and there would be more visual search, which could decrease the DPS. The *leading* cues do not have these problems since they leave some of the centre of view clear and are not lost in the periphery. The *Halo* and *Wedge* cues give constant feedback about the direction to the target, and the user does not have to process a lot. This may be the reason these cues had a high DPS - the user instantly knew which direction to go. *Directing* cues projected space into a cue, which was done in different ways.

Projection mapping

Figure 3 shows that some *directing* cues did well, such as the EyeSee360 with a DPS of 25, but there was also other ones that had a lower DPS, such as the *sidebARs* and *MirrorBall* [1] with the DPS values around 15. The *sidebARs* cue did similar projection as the EyeSee360 by projecting a 3D space into a 2D map, but only had the cues in the right or left side of the HoloLens display. This meant that if a target was above the user, the cue was still at the left or right top corner. The *MirrorBall* cue was like a mirror sphere placed in front of the user that reflected the targets in the mirror. Since the user both had to analyse the real world and the inverted *MirrorBall* at the same time, some confusion could arise. These results indicate

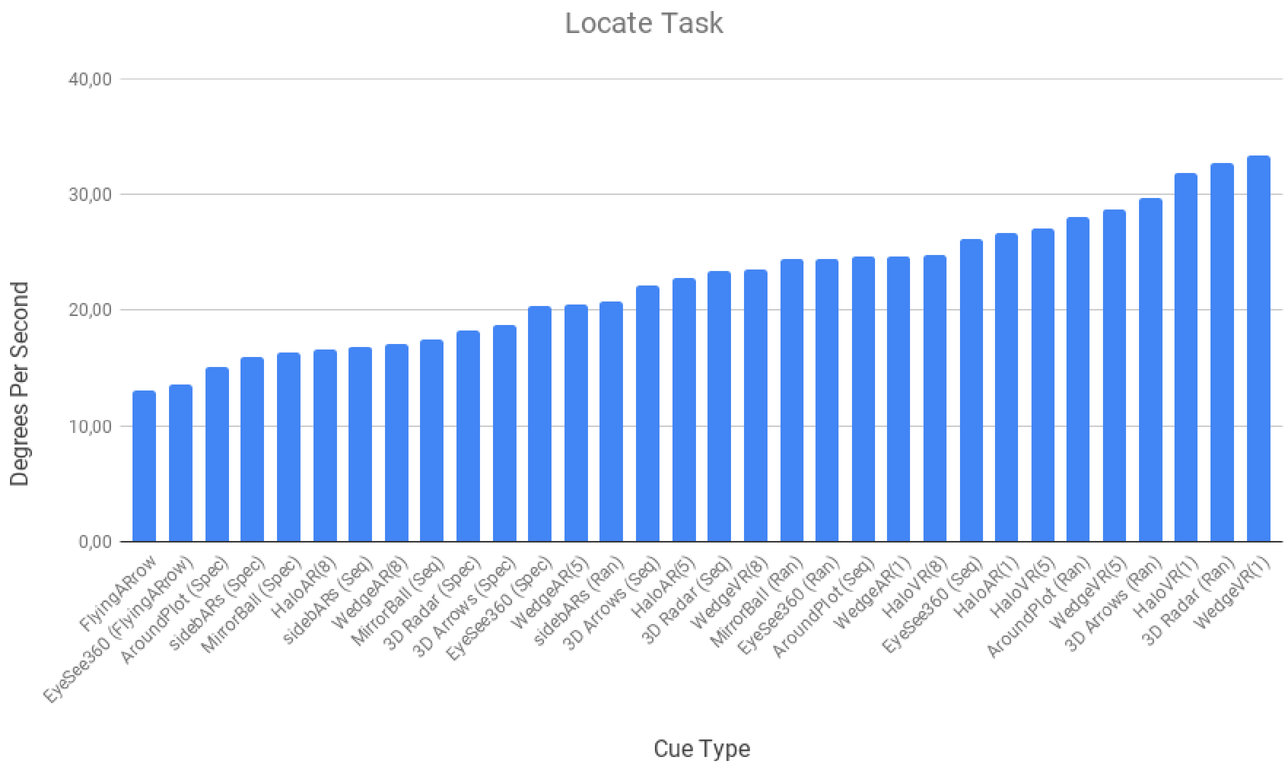


Figure 3: Diagram of degrees per second for 2 DoF Task

that a uncommon projection mapping may be too complex for the user.

Arrows

The studies also used cues in the form of arrows pointing towards a target. In both the Bork et al. and Gruenefeld et al. studies, the arrow cues had a lower DPS than the EyeSee360 cue - 15% slower in Bork et al. [1] and around 4% slower in Gruenefeld et al. [6]. In Bork et al. the arrow cue was always visible in front of the user. Their video demonstration indicates that the arrow's pivot point is not in the arrow it self, but on an invisible sphere that the arrow(s) is put on. In this kind of endogenous arrow cue, the user had to process the direction of the arrow, and then move towards the target. In Gruenefeld et al., the arrow cue, *FlyingARrow* was entirely different. The arrow appeared far away in the centre of view of the display and *moved* physically in a line towards the out-of-view target. This kind of cue worked like the concept of a person in front the user throwing a ball towards an out-of-view target - based on the path of the cue, the target would roughly know the placement of the out-of-view object. It is difficult to compare the DPS results from the Gruenefeld et al. study [6] with the rest of the DPS values since their two DPS values from *FlyingARrow* and *EyeSee360* are unusually low. A reason could be the size of the out-of-view targets used in their test - the smaller the size, the longer the acquisition time would be, but since they did not share their target size it is only a guess. The results from Bork et al. and Gruenefeld et al. indicate that their arrow cues performed slower than their other *directional* cues.

Analysing 3 DoF Tasks

A couple of studies have analysed the use of cues in 3 DoF VR and MR tasks [8, 9, 10]. In these types of tasks, the user had to locate and retrieve an out-of-view target in order to assemble a birdhouse [9], place LEGO bricks in a specific shelf [10] or LEGO structure [8]. The *arrow* and circular cue "*SWAVE*" (Spherical Wave) cues had the highest DPS values, which were around 20-21 DPS, see Figure 4. An interesting observation is that the SWAVE cues, a close relative to the *leading Halo* cues, had roughly the same DPS in both a birdhouse assembly task [9] and the LEGO placement task [10], 3% difference. On the other hand, the *arrow* cue had a higher DPS in the birdhouse task (21), and lower one in the LEGO placement task (16) which makes a 20% difference. A reason for the difference could be that in the birdhouse task, the participant had to sit down for the whole task, whereas in the LEGO task the participant had to stand and walk in order to complete the task. This indicates that in a 3 DoF task where the user does not have to move physically, but only rotate, an arrow cue would convey direction well. On the other hand, in a 3 DoF task where the user has to move in 3D space, a circular cue such as the *SWAVE* would convey direction and distance more effectively than an arrow cue. The design of the arrow cues used in the birdhouse and LEGO studies were 2D arrows - this may also be the reason why it is suited for rotating the user, rather than conveying depth.

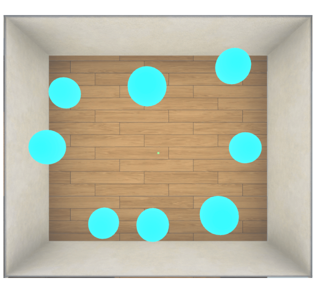
The third 3 DoF task where the user had to place a LEGO brick on a LEGO structure involved placement occlusion [8]. The user had to move around the structure in order to find the target. The *3D path* cue continuously *lead* the user to the target by showing a path from the central FoV to the target. It had a 15% higher DPS than the circular *SWAVE* cue. These results indicate that a *leading* cue with a 3D aspect guides the user to an occluded area more effectively than a *leading* cue that is 2D.

Mental Demand and Preference in 2 Dof and 3 DoF

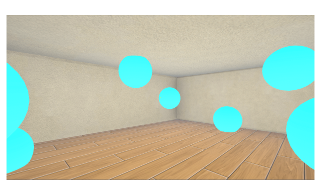
DPS values tell us how fast users move towards an out-of-view target, but they do not give us an idea about mental load or subjective aspects from the users. In the LEGO placement task, the *leading* SWAVE cue had a 28% lower NASA-TLX score than the *directing* arrow cue, 20 and 28 reespectively [10]. In the 2 DoF locating task, the circular cue *HaloAR* had a 8-12% lower NASA-TLX score (21) than the *WedgeVR and WedgeAR* cues (24,25) [7]. These results indicate that circular cues might result in lower mental demand than other types of cues in both 2 DoF and 3 DoF tasks. A reason could be the *amodal completion* aspect in a circular cue makes the mental calculation to a target easy. On the other hand, results indicate that a *directing* cue that relies on projection mapping results in higher mental demand than another *directing* cue that does not rely on projection mapping. The *EyeSee360* (46) had a 18% higher NASA-TLX score than the *FlyingArrow* (39) in a locating task [6]

FIRST STUDY

The goal of this study was to test different cues in a simple room in a 2 DoF task to find out what cue would be the best in that scenario. The participants performed two tasks, a simple and a complex. The simple task only had one target visible at a time. The complex task had all targets visible at the same time, but only one active. The participant was located in the center of the room with the targets around them, see Figure 5a for a top down view of the room.



(a) From above with all the targets visible



(b) From the side with all the targets visible.

Figure 5: Simple Room, the participant is situated in the middle of the room

Method

Task

The first study consisted of a 2 DoF test where the participants had to locate 8 out-of-view targets using three different cues, *EyeSee360*, *Halo* and *FlyingARrow* and one condition without a cue. The total amount of conditions was seven with three repetitions each. After each condition the participants answered the NASA-TLX questionnaire. Four of the conditions had one target visible at a time. The participant performed the test using: *EyeSee360*, *Halo*, *FlyingARrow*, and *No Cue*. In the remaining three conditions the participants had to locate the targets using the same three cues with all the targets visible at the same time. We did not have the participants locate the targets without a cue as it would be very difficult with all targets visible. After the participants completed the conditions with all targets visible they answered a preference questionnaire. The same was done after they finished the last conditions with one target visible.

Measures

The DPS was calculated using the angle between a vector from the centre of the HMD's visual port to a vector going from the participant to the next target. This angle was logged at the onset of a new target and then divided by the acquisition time for acquiring the target, to get the DPS. After each condition, the participant answered a NASA-TLX questionnaire. After all three conditions, the participant answered a 5-point preference questionnaire about the three cues. The participant then vocally stated their opinion of each cues.

Hypotheses

In this study, we sought to test the following hypotheses:

- H_1: *The leading cue Halo will have a higher DPS than the directing cues*
- H_2: *The directing cues will have a higher workload score than the leading cue*
- H_3: *The all targets visible task will have a lower DPS performance compared to the one target task*

EyeSee360

The *EyeSee360* cue is a directing visualisation technique that maps 3D space including target positions to a 2D plane in front of the participant. As mentioned previously, it was developed by Gruenefeld et al. [5] and has been used in other studies [6, 1]. The centre rectangle represents the FoV of the user and each of the dotted lines represents a 45° angle from the user's focus. The vertical lines represents the angle of the horizontal direction and the horizontal lines represents the angle of the vertical direction. The cue leads the user by representing the target with a red dot. It can be seen in Figures 6a & 6b.

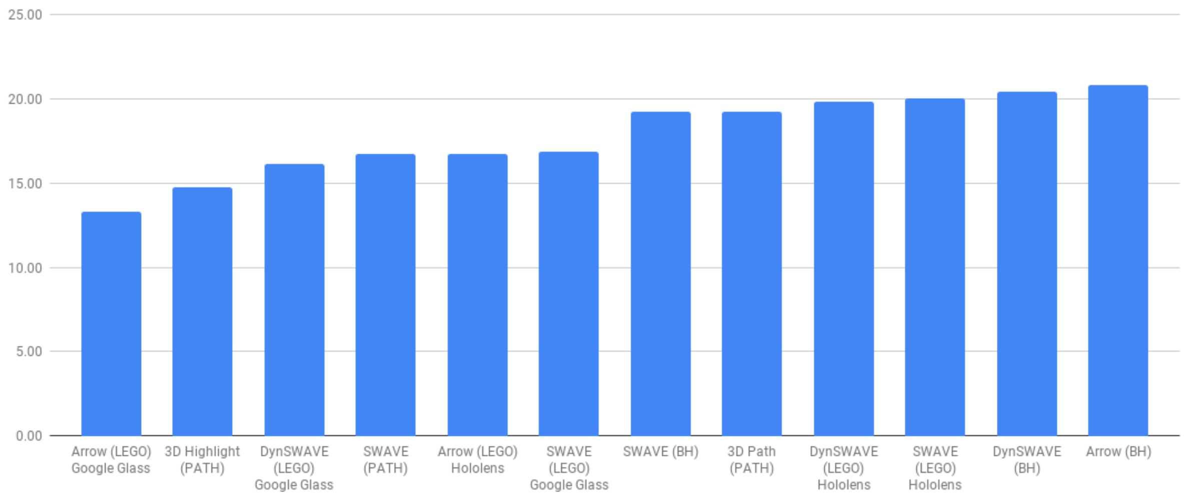
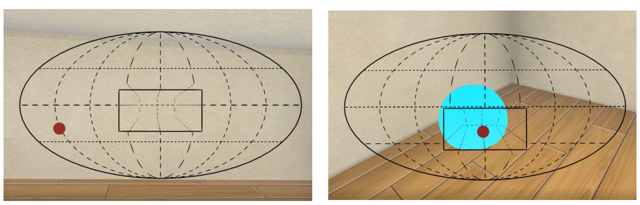


Figure 4: Diagram of DPS values for Assembly Task. BH = Birdhouse[9], LEGO shelf task[10], PATH - LEGO structure task[8]

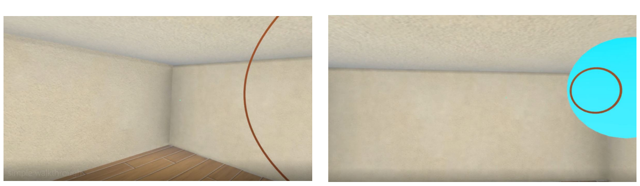


(a) Target out-of-view (b) Target in-view

Figure 6: Stages of the *EyeSee360* cue.

Halo

As one of the most used visual guidance techniques, the *Halo* cue has been implemented in a variety of different ways [**?**, 2]. In our case, we decided to implement a 2D halo that moves in 3D space in relation to the user's orientation. It is a leading cue that guides the user to the target by originating the *Halo* centre at the target and displaying the edge of the circle within the user's FoV, as can be seen in Figures 7a & 7b.



(a) Target out-of-view (b) Target in-view

Figure 7: Stages of the *Halo* cue.

FlyingARrow

The arrow is another one of the more common visual guidance techniques. It also had different variations across studies [6, 7, 9, 2]. In the end, we chose to go with Gruenefeld et al's *FlyingARrow* [6]. The *FlyingARrow* is a 3D guidance technique that directs the user towards a target by having an arrow fly from the centre of the user's view to the out-of-view target. We decided to have the arrow fly multiple times, instead of the single time Gruenefeld et al. did. Additionally, we decided against using the accompanying sound, as we thought that it might give the *FlyingARrow* an advantage over the other cues.

(a) Target out-of-view (b) Target in-view

Figure 8: Stages of the *FlyingARrow* cue.

Results

The following section introduces the results gathered in the first study. In total six people participated in the test (all male) with an average age of 25.

Search Time

The time factors in this study were Degrees Per Second(DPS), and Acquisition Time. These were tested across the conditions for normality with a Shapiro-Wilk test and then for significance with either an ANOVA or a Friedman's test. The DPS for each cue in both tasks can be seen in Figures 9 and 10.

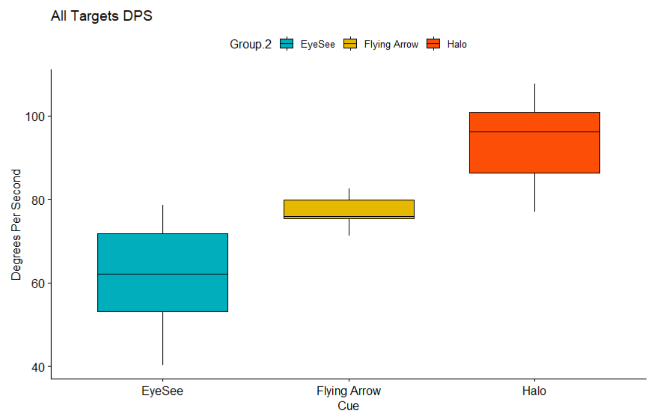


Figure 9: DPS for all targets visible across the three cues.

A Shapiro Wilk test showed normally distributed data in both cases. An ANOVA showed a significant difference in the tasks with all targets visible where the *Halo* cue was significantly faster than *FlyingARrow* and *EyeSee360*. No significant differences were found in the task with one target visible.

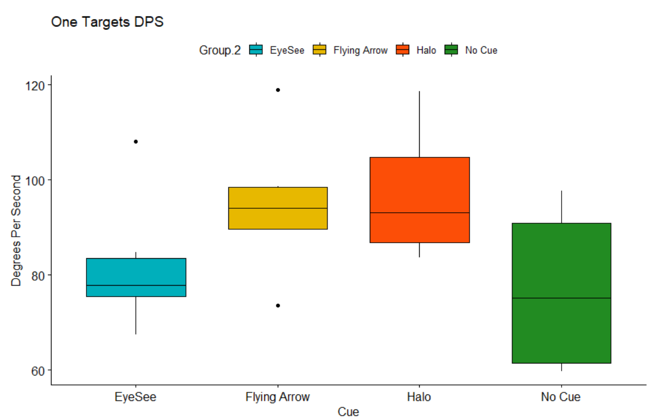


Figure 10: DPS for one target visible across the three cues.

To test if the all targets visible task was indeed a more complicated task than the one target at a time task, the "No Cue" data was removed from the one target task, to balance the data. If a Shapiro-Wilk test showed normal distribution, a paired t-test was used, and if there was no normality, a Wilcoxon signed rank test was used. When comparing the two conditions' DPS, a significant difference was found ($p < 0.001$). The task with all targets visible had significantly lower DPS than the task with one target at a time. Additionally, when looking at acquisition time, the all targets visible task also performed significantly worse than with one target.

NASA-TLX

These answers were tested for significance with a Friedman's test. No significant differences were found. However, a trend can be seen in Figure 11 where *Halo* and *EyeSee360* found to be less mentally demanding than *No Cue* and *FlyingARrow*.

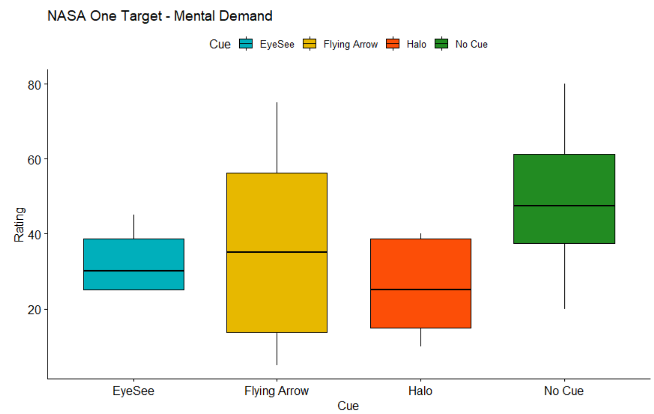


Figure 11: Mental Demand for the cues with one target visible.

Preference

The results can be seen in Figures 22 and 13. People preferred the *Halo* cue in both the all targets visible task as well as the one target visible, but according to a Friedman's test no significant differences were found.

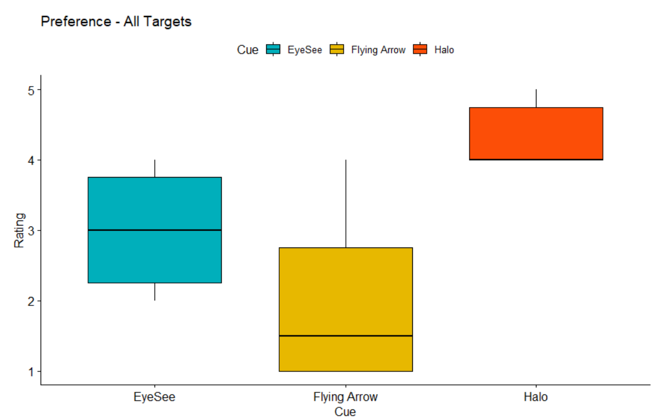


Figure 12: Preference of cues with all targets visible

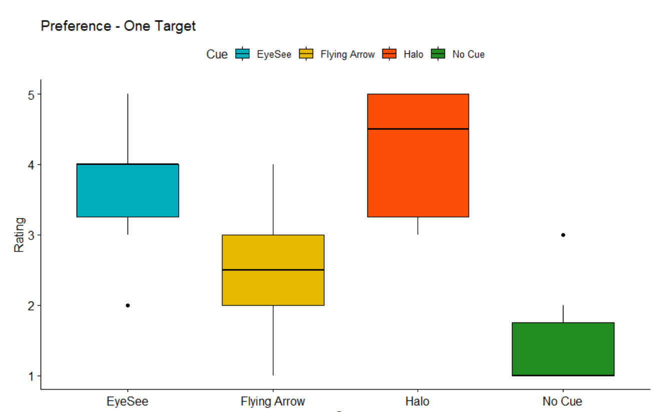


Figure 13: Preference of cues with one target visible

Discussion

Overall the study shows that *Halo* performs well in both tasks. The participants rated it higher than the other cues in preference and also gave it a low score regarding mental demand. The *FlyingARrow* did generally not perform very well. Some participants did not like that the arrow would fly through them sometimes if the targets were behind them. *EyeSee360* did not perform well regarding acquisition time. Due to these results the *EyeSee360* and *FlyingARrow* were not included in the second study. The DPS values we got from the test were much larger than any previous work. This might be because our targets were larger than the targets from the other studies, which would results in faster acquisition time since it was easier to find and "hit".

SECOND STUDY

The goal of this study was to test visual cues in a 3 DoF assembly task. Through our collaboration with a company, we received a VR Satellite assembly project that we modified.

Method

Task

Each participant had to perform 40 small interactive tasks in order to assemble a satellite. In total, the satellite were assembled three times per cue. There were different kinds of physical tasks that the participant did: pick up an item, place the item in the correct place, turn an item, screw bolts or nuts, or press a button. Items were spread out over a big area, around 4x4m, see Figure 14.

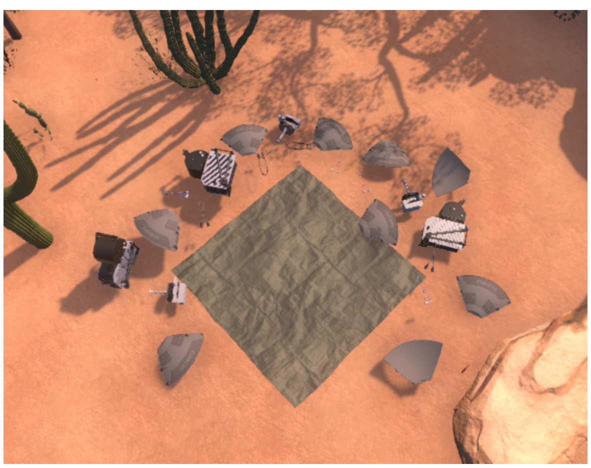


Figure 14: The complex environment seen from above with all the targets visible. The participant was situated in the middle of the tarp.

The environment had duplicate parts of the satellite in order to have ambiguity and visual noise. The participant did not assemble the same satellite in all three conditions, there were actually two. This means that in condition one and three, the participant assembled satellite A, and in condition two the participant assembled satellite B. The difference between the two satellites was only the placement of the parts. The type of tasks were either out-of-view or in-view, e.g. get an out-of-view object, or screw an in-view bolt. This increased the complexity of the whole task. There was always a visual outline that showed which part the user had to interact with, a blue pulsating outline. The satellite assembling would be much more difficult without the outline, especially in some of the detailed tasks. The experiment took around 20-25 minutes.

Measures

After each condition, the participant answered a NASA-TLX questionnaire. After all three conditions, the participants were asked to rate the three cues from 1-5 according to preference with 5 being most preferred. The participant then vocally stated likes and dislikes of the cues. There were three acquisition times logged, the time it took to: have target in-view, touch target, finish task. The in-view time was logged to get the time from out-of-view to in-view. In the tasks where the target was already in-view, the in-view time would be 0, which is why we added the time it took to touch the target. The finish task time was not important because there were many complications with the tasks - e.g. problems with putting the target in the right place, or screwing a bolt. The DPS was calculated in the same way as in the first study, but instead of using acquisition time, it was calculated using "touch time".

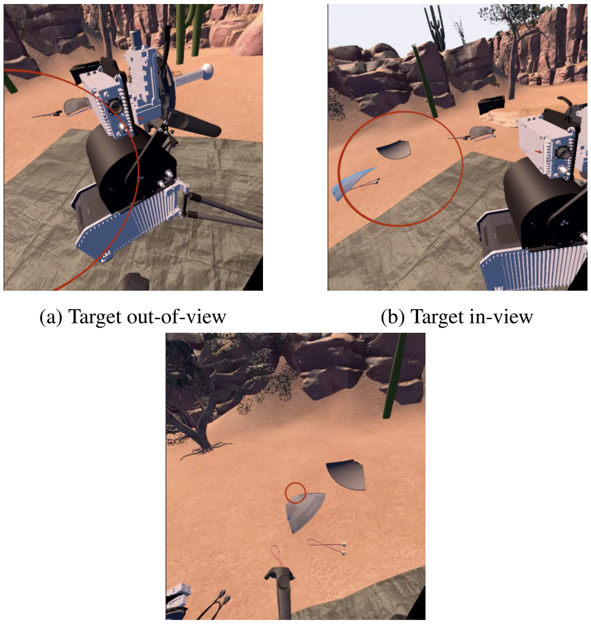
Hypotheses

In this study we sought to test the following hypotheses:

- H_4: *The leading cues Halo and Wedge will outperform the directing cue FIV Arrow in regards to DPS*
- H_5: *The leading cues Halo and Wedge will be less workload intensive than the directing cue FIV Arrow*
- H_6: *Participants will prefer the more discrete FIV Arrow cue over the Halo and Wedge cues*

Halo

The *Halo* used in this study was based on the same implementation as the first study. A difference from the first study, was that it would not vanish when the target was in view, but rather shrink to a defined minimum size, as seen in 15c.



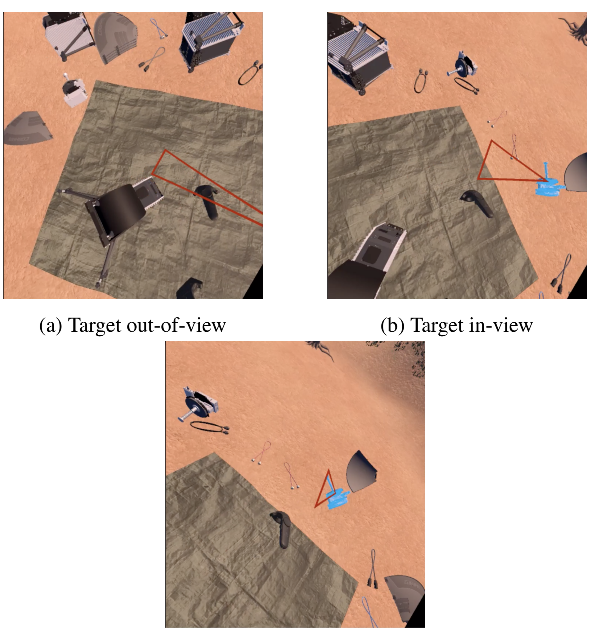
(a) Target out-of-view (b) Target in-view

(c) Target in centre view

Figure 15: Stages of the *Halo* cue

Wedge

The *Wedge* is the final one of the most common visual guidance techniques. Similar to the *Halo* technique, we decided to use it as a 2D *Wedge* that moves in 3D space. The *Wedge* is another leading cue that guides the user to the target by having the tip of the *Wedge* originate at the target and displaying the other end within the user's FoV, as seen in Figure 16.



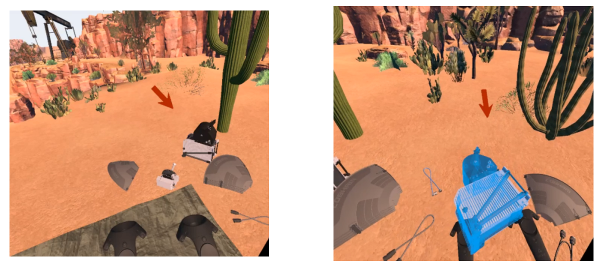
(a) Target out-of-view (b) Target in-view

(c) Target in centre view

Figure 16: Stages of the *Wedge* cue

FIV Arrow

The findings from our first study showed that the FlyingARrow cue did well in regards to DPS performance, but it was far from the most preferred one. Based on these findings, we decided to include a new variation of the arrow cue, seen in Figure 17. As the FlyingARrow was described as intrusive, we decided to have it be a 3D arrow Fixed-In-View (FIV) rather than having it move towards the target continuously.



(a) Target out-of-view (b) Target in-view

Figure 17: Stages of the *FIV Arrow* cue

Results

The following section introduces the results gathered during the second study. 22 people participated in the study (19 male, 3 female), with a mean age of 23.

Search Time

The time factors were Degrees Per Second(DPS), In-View Time, Touch Time, and Total Acquisition Time. DPS is calculated using Touch Time.

With DPS, a Shapiro-Wilk test showed the data was normally distributed, which led to performing an ANOVA to test for significance. The test showed a p-value above 0.05, so no significance was found.

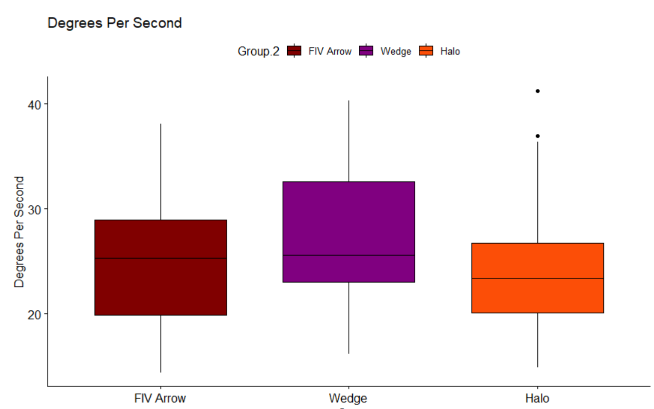


Figure 18: DPS for the three cues.

Figure 19 shows a box plot of the results from the in-view time. The in-view time refers to the time it takes the participant to get the object within a FoV of 40°. It can be seen that the *Wedge* cue had the fastest in-view time of the three cues with a median time of 1.135s.

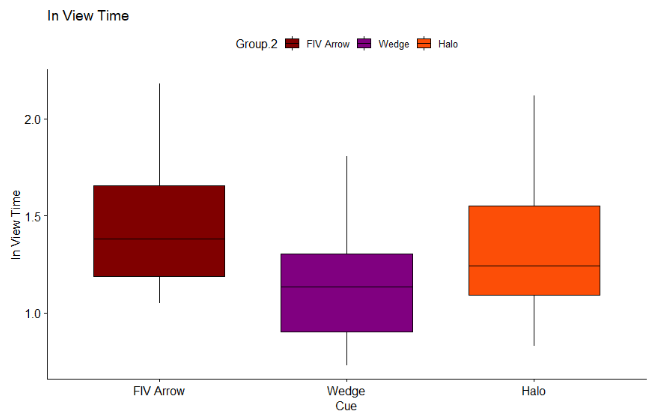


Figure 19: In-View time for the three cues.

A Shapiro-Wilk test showed that the data was not normally distributed ($p < 0.01$). Therefore, a Friedmans test followed by a post-hoc Wilcoxon signed rank test with Bonferroni correction was used and showed a significant difference with *Wedge* being faster than *Halo* ($p < 0.01$) and *FIV Arrow* ($p < 0.00001$).

Both touch time and acquisition time were shown to be normally distributed by a Shapiro-Wilk test, and the following ANOVAs showed no significant differences.

NASA-TLX

The results of each score from the NASA-TLX questionnaire was analysed using a Friedman's test, which was then followed up, if significant, with a post-hoc Wilcoxon signed rank test with Bonferroni correction. The scores with significant differences are mentioned below.

The *Halo* cue scored a higher rating in frustration than *Wedge* and *FIV Arrow* in the NASA-TLX questionnaire, as can be seen in Figure 20. Meaning the participants on average found *Halo* to be significantly more frustrating.

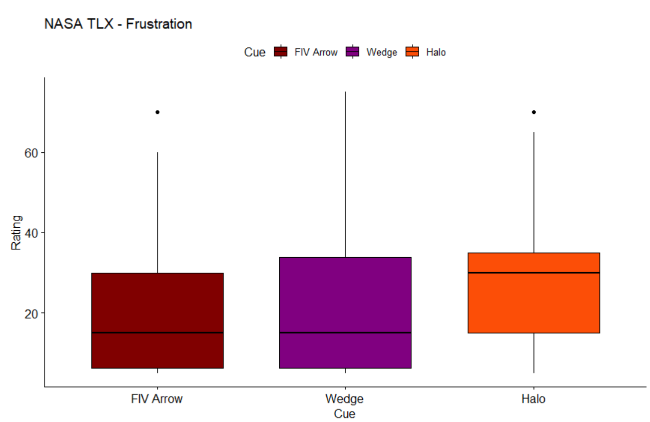


Figure 20: Frustration rating from NASA-TLX questionnaire. Lower is better.

The same can be said for mental demand regarding the *Halo* cue, as seen in Figure 21. The cue also scored higher from the participants in that regard. After the tests, the participants were asked about their thoughts on the visual cues. Ten participants expressed negative thoughts on the *Halo* cue, such as it being too big and obtrusive, and difficult to know the direction of the target.

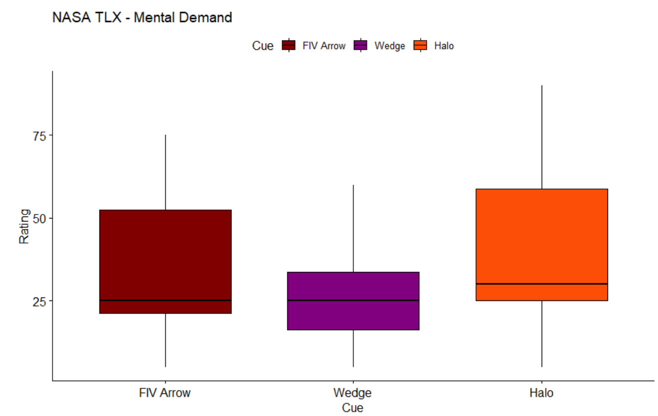


Figure 21: NASA-TLX score for mental demand. Lower is better.

Preference

The ratings show that the *FIV Arrow* was the most preferred and the *Halo* was the least preferred cue, however no significant differences were found from a Friedman's test.

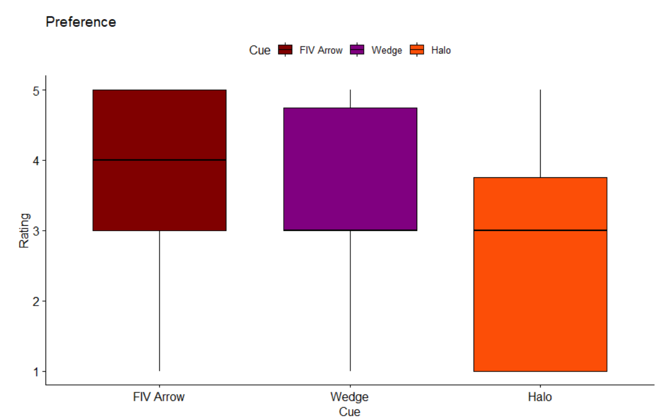


Figure 22: Preference score from 1 (Not preferred) to 5 (Very much preferred).

DISCUSSION

From the results we can see that *Halo* did not perform as well as expected from the results in the first study. Based on preference score and verbal feedback from the participants we saw a trend that the participants did not like the *Halo* because of the amount of space it took up in their FOV. The mental demand and frustration NASA-TLX score for *Halo* in our first study was lower than in the second study, which indicates that the participants preferred a less obtrusive cue when performing a 3 DoF task. This was also supported by the verbal feedback. Participants mentioned that the *Halo* was not good for the in-view tasks and that it worked better at leading towards out-of-view targets. This was also mentioned for the *Wedge* cue. Some participants preferred the *FIV Arrow* cue for in-view tasks together with the highlight.

The objects the participants had to find in the second study were always on the ground, so the angle with which the DPS was calculated was affected by this.

Across the two studies, six hypotheses were set up. We do not accept $\mathbf{H}_1$: *Leading* cues did have a higher DPS than the *directing* cues in the "all targets visible" locate task, but not in the "one target visible" task. We do not accept $\mathbf{H}_2$: *Directing* cues did not have any significantly higher NASA-TLX scores than the *leading* cues in the locate task. We accept $\mathbf{H}_3$: The all targets visible locate task did have a significantly lower DPS performance compared to a one target visible task. We reject $\mathbf{H}_4$: The *leading* cues *Halo* and *Wedge* did not outperform the directing cue *FIV Arrow* in regards to DPS. We reject $\mathbf{H}_5$: The *leading* cues *Halo* and *Wedge* did not receive any significantly lower NASA-TLX scores compared to the *directing* cue *FIV Arrow*. However, the *Halo* had a significantly higher "frustration" value compared to the other two cues and higher "mental demand" value than the *Wedge*. We reject $\mathbf{H}_6$: We found no significant differences in preference between the *directing* cue *FIV Arrow* compared to the *leading* cues *Halo* and *Wedge*.

Even though counterbalancing was used in the tests, a learning curve could not be avoided completely, as the participants quickly learned the order of the tasks and where to find the objects in spite of them having different locations.

CONCLUSION

Initially we asked the question: *How do cues guide the users in tasks with differing complexity and which cues are best in those tasks?* Here we conclude upon this.

In the locating task, the *Halo* cue outperformed the *Flying Arrow* and *EyeSee360* cue in both conditions. It had significantly faster DPS than the other two when all targets were visible and it also outperformed the others in both the NASA-TLX and preference, although not significantly. These results suggests that the *leading* cue *Halo* is the best cue for a locating task.

In the assembly task, the *leading* cue *Wedge* cue performed best in regards to time and DPS. It had significantly faster In-View Time compared to both *FIV Arrow* and *Halo*. *Wedge* also had significantly less mental demand than the *Halo* cue, as well as both *Wedge* and *FIV Arrow* being significantly less frustrating than *Halo*. In regards to preference, both *Wedge* and *FIV Arrow* scored higher than *Halo*, though not significantly.

REFERENCES

1. F Bork, C Schnelzer, U Eck, and N Navab. 2018. Towards Efficient Visual Guidance in Limited Field-of-View Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2983–2992. DOI: http://dx.doi.org/10.1109/TVCG.2018.2868584
2. Stefano Burigat, Luca Chittaro, and Silvia Gabrielli. 2006. Visualizing locations of off-screen objects on mobile devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services - MobileHCI '06*. ACM Press, New York, New York, USA, 239. DOI: http://dx.doi.org/10.1145/1152215.1152266
3. Marisa Carrasco. 2011. Visual attention: The past 25 years. *Vision Research* 51, 13 (jul 2011), 1484–1525. DOI:http://dx.doi.org/10.1016/J.VISRES.2011.04.012
4. John M Carroll. 2003. *HCI Models, Theories, and Frameworks : Toward a Multidisciplinary Science*. Elsevier Science & Technology, San Francisco, UNITED STATES. http://ebookcentral.proquest.com/lib/aalborguniv-ebooks/detail.action?docID=294610
5. Uwe Gruenefeld, Dag Ennenga, Abdallah El Ali, Wilko Heuten, and Susanne Boll. 2017. EyeSee360: Designing a Visualization Technique for Out-of-view Objects in Head-mounted Augmented Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction - SUI '17*. ACM Press, New York, New York, USA, 109–118. DOI: http://dx.doi.org/10.1145/3131277.3132175
6. Uwe Gruenefeld, Daniel Lange, Lasse Hammer, Susanne Boll, and Wilko Heuten. 2018. FlyingARrow: Pointing Towards Out-of-View Objects on Augmented Reality Devices. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays (PerDis '18)*. ACM, New York, NY, USA, 20:1—-20:6. DOI: http://dx.doi.org/10.1145/3205873.3205881
7. Uwe Grünefeld, Abdallah El Ali, Susanne Boll, and Wilko Heuten. 2018. *Beyond Halo and Wedge: Visualizing Out-of-View Objects on Head-mounted Virtual and Augmented Reality Devices*. DOI: http://dx.doi.org/10.1145/3229434.3229438
8. P Renner, J Blattgerste, and T Pfeiffer. 2018. A Path-Based Attention Guiding Technique for Assembly Environments with Target Occlusions. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 671–672. DOI: http://dx.doi.org/10.1109/VR.2018.8446127
9. Patrick Renner and Thies Pfeiffer. 2017a. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*. IEEE, 186–194. DOI:http://dx.doi.org/10.1109/3DUI.2017.7893338
10. P Renner and T Pfeiffer. 2017b. [POSTER] Augmented Reality Assistance in the Central Field-of-View Outperforms Peripheral Displays for Order Picking: Results from a Virtual Reality Simulation Study. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. 176–181. DOI: http://dx.doi.org/10.1109/ISMAR-Adjunct.2017.59
11. Colin Ware. 2004. *Information Visualization : Perception for Design*. Elsevier Science & Technology. http://ebookcentral.proquest.com/lib/aalborguniv-ebooks/detail.action?docID=226683