Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display
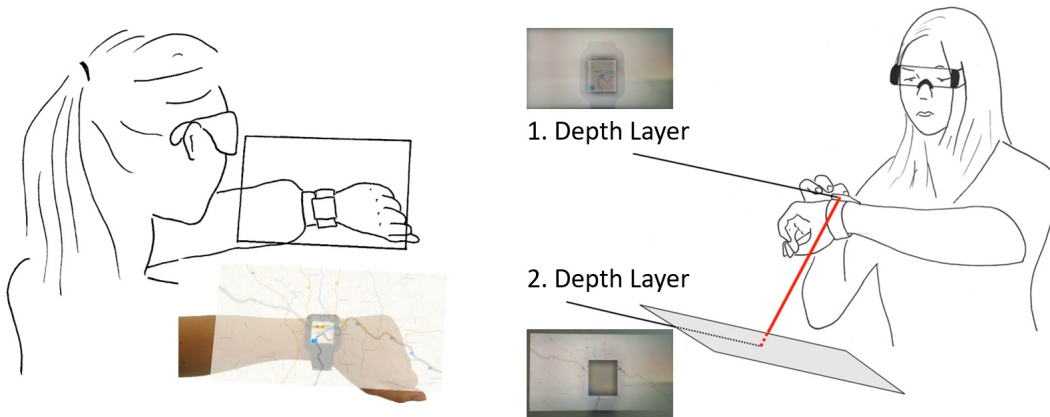
Anna Eiberger
University of Passau
eiberg01@stud.uni-passau.de

Per Ola Kristensson
Cambridge University
pok21@cam.ac.uk

Susanne Mayr
University of Passau
susanne.mayr@uni-passau.de

Matthias Kranz
University of Passau
matthias.kranz@uni-passau.de

Jens Grubert
Coburg University of Applied
Sciences and Arts
jens.grubert@hs-coburg.de

Figure 1: Left: Vision of extending the field of view of body-proximate displays such as a smartwatch through an optical see-through head-mounted display with visual information presented at a uniform depth layer. Right: Visualization of the two depth layers (smartwatch and focal distance of the head-mounted display), which leads to the need of accomodation and vergence changes to integrate information accross both displays.

ABSTRACT

Optical see-through head-mounted displays (OST HMDs) typically display virtual content at a fixed focal distance while users need to integrate this information with real-world information at different depth layers. This problem is pronounced in body-proximate multi-display systems, such as when an OST HMD is combined with a smartphone or smartwatch. While such joint systems open up a new design space, they also reduce users' ability to integrate visual information. We quantify this cost by presenting the results of an experiment (n=24) that evaluates human performance in a visual search task across an OST HMD and a body-proximate display at 30 cm. The results reveal that task completion time increases significantly by approximately 50% and the error rate increases significantly by approximately 100% compared to visual search on a single depth layer. These results highlight a design trade-off when designing joint OST HMD-body proximate display systems.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

augmented reality, perception, multi-display environments, accomodation, vergence

1 INTRODUCTION

Optical see-through head-mounted displays (OST HMDs) are on the verge of becoming mainstream computing platforms, and certain devices, such as the Microsoft HoloLens, are already in-use in various industries, such as construction engineering.

Body-proximate displays, such as smartwatches and smartphones have the potential to augment OST HMDs [17, 33] and provide two primary benefits. First, they can extend the available display space and thereby provide new opportunities for distributing information to the user (see also Figure 1, left). Second, they can introduce a secondary capacitive touchscreen input channel to the OST HMD system. As pure OST HMD input is inherently noisy and error-prone due to a reliance on voice, eye-tracking, head-tracking or gesture recognition, such a secondary input channel is valuable as it provides relatively certain low-noise information about users' intentions. For instance, a joint OST HMD-smartwatch system was recently studied with respect to indirect control of the cursor on the OST-HMD via the smartwatch [50].

Other potential applications include distributing notifications or warnings across a phone and an OST HMD, showing context-relevant information about the visual scene depicted in the OST HMD on the phone, or augmenting OST HMD interaction with a physical input/output device, for instance, by allowing users to move their phone in front of their face and observe a dynamic lens of the visual scene which can be modulated by the user interacting directly with the phone.

However, while body-proximate displays may provide these and possibly other benefits to OST HMD systems, the implications of using such joint systems are currently underexplored. In general, while distributing information across a body-proximate display and an OST HMD can open up new user interface solutions, such distribution of information is likely to negatively affect users' ability to integrate information (see Figure 1, right). Hence, it is neccessary to quantify the cost of jointly processing visual information across OST HMDs and body-proximate displays.

With the exception of Magic Leap Creator's Edition, commercial OST HMDs present virtual information at a single focus distance. Vari-focal HMDs (e.g., [1, 13, 48]) are also still not widely commercially available.

Prior work [15, 24] indicates that if information in the physical scene has to be integrated with virtual information, the user's visual and cognitive load increases. However, this prior research has mostly focused on users integrating virtual information on an OST HMD display with information in the physical scene in the medium or far visual field, such as visual search tasks or depth perception judgment tasks. Work focusing on OST HMD information integration in the near field are rare, with a few exceptions, such as work investigating depth-perception [43] and object localization [14].

In this paper, we complement prior research efforts by investigating the joint performance of an OST HMD coupled with a body-proximate display. We report of an experiment comparing reaction time, error rate, workload and degree of simulator sickness induced in a visual search task carried out either solely in the depth field of the OST HMD or distributed across the depth field of the OST HMD and the body proximate display. We find that the distribution of the visual search task results in a significant

(approximately 50%) increase in reaction time and a significant (approximately 100%) increase in error rate as well as in increases in workload, however, no significant difference in perceived simulator sickness. These results quantify the substantial information integration cost a joint OST HMD-body-proximate system incurs for users and suggest that there is a significant trade-off when using joint OST HMD and body-proximate display systems.

2 RELATED WORK

The work builds on a large body of related work from the domains of perceptual psychology, augmented reality and multi-display systems. In the following, we highlight work from the domains body-proximate multi-display systems and perceptual issues in OST HMD.

2.1 Perceptual Issues in OST HMDs

Commercially available binocular OST HMDs display virtual content typically at a single focal distance. Various perceptual issues when using OST HMDs have been explored (c.f. [28] for an overview). Specifically, if content from the physical world at a different focal distance needs to be integrated with the virtual content, this can lead to the need of repeated accommodation changes as well as context switches, resulting in increasing visual and cognitive attention [15]. Relevant factors are differences in focal distance and vergence as well as the visual angle between physical and virtual content.

Depending on the expected distance of content in the physical world, different focal distances have been proposed for Head-up displays (e.g., 2m or optical infinity) with no concluding recommendations, yet [47, 51]. Gabbard et al. highlighted the need for multi-focal Augmented Reality (AR) displays [16] with several research protoypes beginning to emerge [1, 13, 48]. For current single focus HMDs, Oshima et al. [34] and Cook et al. [10] proposed and evaluated a system for adaptive sharpening HMD display content. Due to commercial availability, our study used a collimated off-the shelf OST HMD.

There have been a number of studies, investigating the effects of processing information across OST HMDs and the physical world. Huckauf et al. [24] investigated the effects of context switching (i.e. switching of visual and cognitive attention between virtual and physical world information) between a monocular OST HMD and a CRT-monitor, both placed at the same focal distance of 61 cm. Their results indicate that there is a cost of context switching ca. 10% between a CRT and an OST-HMD when both displays are at the same focus distance).

Gabbard et al. [15] also investigated the effects context switching, but in conjunction with focal distance switching (i.e. accomodation change). The investigated focal distances were 0.7 m, 2 m and 6 m. Their findings indicate that both focal distance and context switching result in significantly reduced performance in a reading task. Their work shows that both context and focal distance switching are important AR user interface design issues. Our study builds up on this prior work by investigating the joint performance of a (handheld or body-worn) display at typical viewing distance of a smartphone [2] with an OST HMD.
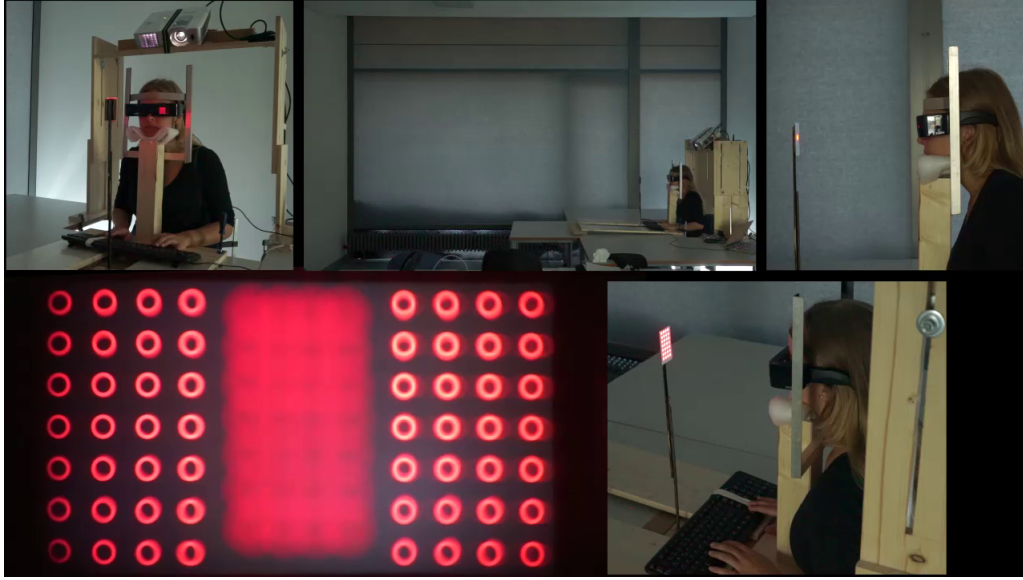
Figure 2: Multiple views on a participant in condition HMD+Projector. Top left: Front view. Top middle: side view of the experiment room. Top right: close-up side view with projection area on the left. Bottom left: view through the HMD with focus on the HMD plane (hence, the center four columns appear blurred)
. Please note, that in the actual first person view of the user, she would see double images of the plane out of focus due to stereopsis). Bottom right: Back view with projection area.

Winterbottom et al. [49] studied multi-focal AR displays and empirically determined suitable focus distances for in the context of a flight simulator.

Within near-field reaching distances, Singh et al. [43] investigated the effects of focal distance, age and brightness when matching the depth between a real and virtual object using an haploscope within distances between 33.5-50 cm in a depth-matching task [46]. Their results indicate that collimated graphics (i.e. graphics shown at infinite focal distance) resulted in inferior performance compared to graphics shown at a nearer focal distance.

Ellis and Menges [14] studied the effects of vergence, accommodation, observer age, viewing condition and the presence of an occluding surface in the near-field. They found that accuracy is degraded by monocular viewing and an occluding surface. McCandless et al. [42] studied motion parallax and latency in monocular viewing and found reduced accuracy with increasing distance and latency. Singh et al. [44] found that an occluding surface has complex accuracy effects, and Rosa et al. [39] found that accuracy increased when using redundant tactile feedback.

Our study builds on this prior work and extends it to the domain of body-proximate multi-display systems.

2.2 Body-Proximate Multi-Display Environments

Combining mobile body-proximate displays into multi-display environments has been repeatedly investigated. However, the majority of previous works has studied the joint interaction between multiple handheld devices, such as smartphones or tablets, body-worn devices, such as smartwatches or pico-projectors. Only few works dealt with joint interaction between HMDs and smartwatches or smartphones.

Chen et al. studied joint interactions between a smartphone and a smartwatch [8]. Rädle et al. [36] proposed a lamp-mounted spatial tracking system for interacting accross multiple tablets . Grubert et al. [18, 19] extended the idea to not use an instrumented environment, but the user's head as reference point instead.

Houben et al. presented a toolkit for prototyping smartwatch-centric cross-device interactions [23]. A number of different cross-device toolkits have been proposed for focusing on smartwatch, web, or Augmented Reality centric development aspects, e.g., [9, 25, 32, 40, 42, 45].

Grubert et al. [20, 35] presented a survey on challenges in mobile multi-display ecologies and highlighted potential perceptual issues in body-proximate multi-display systems, such as display contiguity [7, 38], visual attention [22, 31, 37] or visual overload [12, 30] which was followed up by a recent survey of Brudy et al. [5].

Regarding the use of OST HMDs in multi-display environments, Serrano et al. [41] investigated the joint use of OST HMDs within a desktop environment. Grubert et al. [17] proposed several interaction techniques between an OST HMD, a smartwatch and smartphone and conducted performance tests of the joint system, but did not compare it against a single device baseline. Budhirja et al. [6] proposed using a smartphone as indirect pointing device for an HMD, an idea late extended unobtrusive input devices [11]. Wolf et al. [50] determined the performance envelope of in-air direct and smartwatch indirect control for OST HMDs.
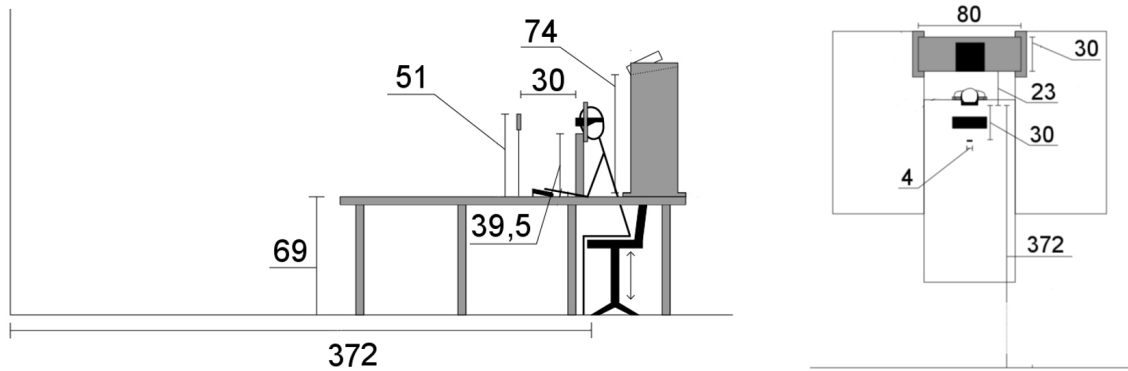
Figure 3: Side view (left) and top view (right) on the apparatus setup. All distances in cm.

Our work informs researchers, designers and developers of joint interactive systems between OST HMDs and body-proximate displays, such as smartwatches and smartphones about the costs of joint information presentation accross those devices.

3 EXPERIMENT

The experimental design is inspired by previous experiments on context switching and focal distance switching [15, 24] and extends those previous investigations to a combination of a OST HMD and a body-proximate display. The primary research question we are addressing is a quantification of how well users are able to integrate visual information jointly presented across an OST HMD and a display at a typical viewing distance of a smartphone [2].

To this end, we asked participants to conduct a visual search task either on an HMD alone or across the HMD and a display at a closer focus distance.

3.1 Design

The experiment used a within-subjects design with one independent variable DEPTHLAYERS. It had two levels: In the HMD+PROJECTOR condition the task was distributed across the HMD (in stereo) and a second screen illuminated by a projector in front of the user. In the HMD condition, the task was conducted at the single HMD screen in stereo mode. We decided against a possible third condition, conducting the task using the projector only at near distance as a pilot study indicated no significant differences between projector only and HMD.

3.2 Participants

Twenty-four volunteers participated in the experiment (12 male, 12 female, mean age 24.7 years, sd = 4.3). All but one participants exhibited normal or corrected to normal vision using corrective glasses or contact lenses. One participant exhibited 20/33 (6/10, logMAR: 0.22) visual acuity with contact lenses but could conduct the study without problems.

3.3 Apparatus and Materials

The apparatus is shown in Figures 3 and 2. A Benq W1070 Full HD projector was used for projection on 4×7.5 cm sheet of cardboard in 30 cm distance of the user. This distance corresponds to a typical smartphone and smartwatch viewing distance [2, 4]. The projector was mounted on a stand above the user's head. The HMD used was an Epson Moverio BT-100 (horizontal field of view: 30.3°vertical field of view: 11.4°). The HMD displayed content at infinite focus and the default vergence distance of 370 cm, which corresponded to a wall at the viewing direction of participants. The field of view of the HMD was too small to set the vergence distance at the same distance as the projector. The head of the user was fixated on a chin rest (two chin rests were used to accommodate for different user heights). Colors, brightness and resolution of the HMD and the projector were matched as close as possible. A node.js server was used to synchronize the image presentation between the projector and the HMD. There was no noticeable delay between projector and HMD image presentation.

The use of other displays, such as a smartwatch or smartphone panel display or other HMDs (as well as resolutions, contrasts, slight variations in distance) likely would lead to variations in the absolute values of the results to be reported. In a pre-experiment, we considered a panel display (Samsung S8+ smartphone) as body-proximate display, which indicated comparable results to the projector. We ultimately chose a projector as the smartphone (curved) bezel would have interfered with the display of the adjacent symbols on the projector. Also, we are not aware of evidence suggesting that the significance or scale of results would substantially change by interchanging a projector display with a panel display.

A physical keyboard, placed on a table at the users hands, was used to get user feedback for both tasks.

For the primary task, a matrix of 84 circular symbols arranged in 7 rows by 12 columns was presented, see Figure 4, left and middle.

The target symbol "0" (number zero) had a width of 0.66° (40′) and height of 1.00° (60′). The distractor symbol "O" (letter O) had a width of 0.85° (51′) and height of 0.94° (56′).

In condition HMD, all symbols were displayed on the HMD. In the condition HMD+PROJECTOR the middle 4 columns were presented on the display at 30 cm distance. The target would show up with a likelihood of 50% and its location on the grid was uniformly distributed in both conditions.

In condition HMD+PROJECTOR, if the target would show up, then it would appear in 33.3% of all cases on the 4 left, 4 middle or 4 right coloumns (i.e. 66.6% on the HMD and in 33.3% on the front
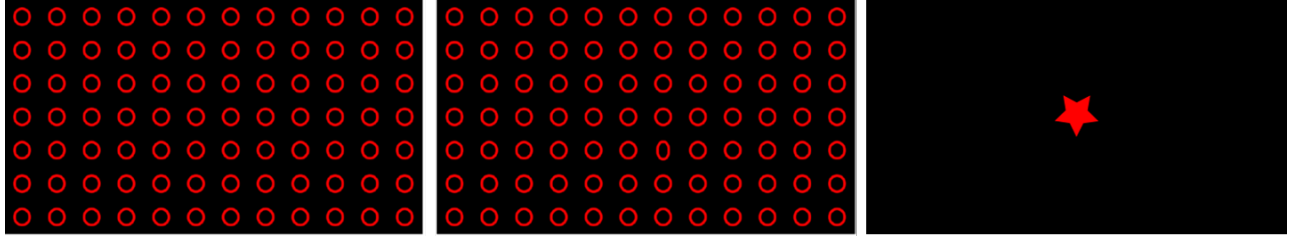
Figure 4: Left: stimulus image with no target present. Middle: stimulus image with target present in row 5, column 7. Right: secondary task with star pointing downwards.
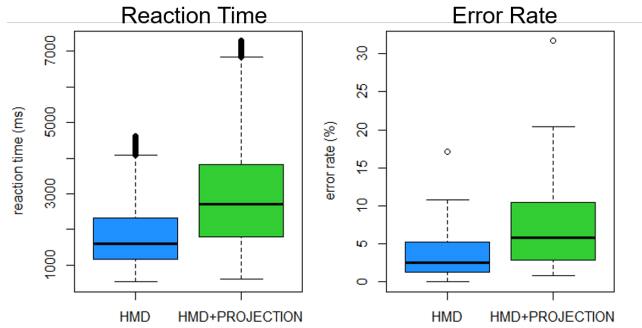

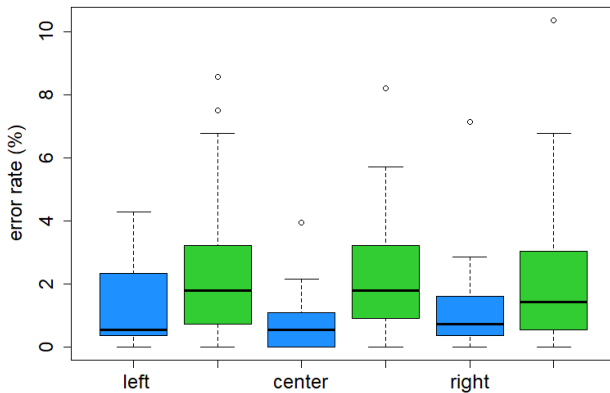**Figure 5: Left: reaction time in ms. Right: error rate.**

Figure 6: Error rate split between left, center and right zones, with each zone encompassing 4 columns. In the condition HMD+PROJECTOR (green) the left and right zone correspond to the columns presented on the HMD and the center zone corresponds to the columns presented on the projector. Condition HMD is depicted in blue. No significant differences where indicated between the zones

projection). The left arrow key should be pressed if the target was present, otherwise the right arrow key. After pressing a key an audio signal would indicate whether the decision was correct or not. The participants had the chance of getting used to the mapping in a trial session.

The secondary task was conducted between images of the primary task, i.e. in alternating order. Here, a star shape would show up, see Figure 4, right. The participant was asked to decide if a star prong was pointing upwards (B) or downwards (space bar). In condition HMD the star was shown on the HMD and in condition HMD+PROJECTOR it was shown on the front projection.

3.4   Task and Procedure

The primary task was to identify whether a target symbol ("0", the number zero) was present among distractors ("O", the letter O). In a secondary, alternating task, participants had to decide whether a star was pointing upwards or downwards. The purpose of the secondary task was solely to erase the previous picture from the primary task from the participant's visual memory. Hence, we exclude the reporting of the results for the secondary task.

Each participant filled out a demographic questionnaire and their visual acuity was determined using the Freiburg visual acuity test [3] to ensure that the displayed symbols could be recognized. All participants were found to be suitable for inclusion into the experiment.

For each condition, the participant was shown 280 separate symbol matrices executed in 4 separate blocks of 70 images each in the primary task, resulting in a total of 560 task executions per participant. The order of the starting condition was counterbalanced across all participants. After each condition the participants filled out the unweighted NASA TLX [21] questionnaire and the Simulator Sickness Questionnaire (SSQ) [26].

The experiment was carried out in a single ca. 60-minute session structured as a 20-minute introduction and briefing phase, a 30-minute testing phase (ca. 10 minutes for condition HMD, ca. 15 minutes for condition HMD+PROJECTOR + five-minute breaks including questionnaires), and 10 minutes for final questionnaires, interview and debriefing. Also, the participants were allowed to rest between individual blocks but did not make use of this option.

3.5   Results

Statistical significance tests for reaction time were carried out using paired, two-tailed t-tests with a significance level of $\alpha = 0.05$. Gender-specific differences for reaction time were tested for using indeptendent t-tests. The analysis were checked for appropriateness against the dataset, log transformed data was used for reaction

time. Subjective feedback and error rate were analyzed with non-parametric Wilcoxon signed-rank tests. Gender-specific differences were tested for using Mann-Whitney U tests.

As the purpose of the secondary task was to erase the previous picture from the primary task from the participant's visual memory, we report on the resulty of the primary task only. Next, we report on reaction time, error rate (i.e. the number of false user decisions divided by the total number of stimuli images), workload (as measured by NASA TLX [21]) and simulator sickness (as measured by the simulator sickness questionaire SSQ [26]). We further report on the search strategies employed by participants.

3.5.1 Reaction Time. The mean reaction time for HMD was 2171 ms (sd = 846) and for HMD+PROJECTOR 3233 ms (sd = 1217), see Figure 5, left. A two-tailed paired sample t-test on log transformed data (to reduce skewness of the data) indicated a significant difference in reaction time ($t(23) = -5.58, p < 0.001, Cohen's\ d = 1.14$). Also, no block-wise or gender-specific differences within conditions for reaction time were found.

In other words, the reaction time using HMD+PROJECTOR significantly increased about 50% compared to only HMD.

3.5.2 Error Rate. The mean error rate for HMD was 3.84% (sd = 3.95) and for HMD+PROJECTOR 8.30 (sd = 7.68), see Figure 5, right. A Wilcoxon signed rank test indicated a significant difference ($Z = -3.16, p = 0.0016, Cohen's\ d = 1.02$). Note that while there is a 50% random chance of being correct, this is equal for both conditions. Also, no block-wise differences within conditions or gender-specific differences for error rate were found.

We further analyzed the error rate according to the left, center and right zones of the displays, with each zone encompassing 4 columns, to check for possible edge effects, i.e. the possibility of a higher error rate in the outer display regions. In the condition HMD+PROJECTOR the left and right zone correspond to the columns presented on the HMD and the center zone corresponds to the columns presented on the projector. No statistically significant differences between the center and the outer zones could be found, see Figure 6.

In other words, the error rate using HMD+PROJECTOR significantly increased about 100% compared to only HMD.

3.5.3 Workload. Scores for the workload of both conditions measured by the unweighted NASA TLX questionnaire are depicted in Figure 7.

Wilcoxon signed rank tests indicated significant differences for mental demand ($Z = -2.61, p = 0.009, Cohen's\ d = 0.81$), temporal demand ($Z = -2.02, p = 0.043, Cohen's\ d = 0.61$), effort ($Z = -3.68, p < 0.001, Cohen's\ d = 1.25$), frustration ($Z = -3.09, p = 0.002, Cohen's\ d = 0.99$) and overall demand ($Z = -3.44, p = 0.001, Cohen's\ d = 1.14$), but not for physical demand or performance. Please note, that a gender-specific difference was indicated for the physical demand rating in the HMD+PROJECTOR condition between female ($mean\ score = 41.67, sd = 28.56$) and male ($mean\ score = 17.92, sd = 20.07$) ($Z = 2.16, p = 0.033, Cohen's\ d = 0.66$), but not within each group between HMD+PROJECTOR and HMD.

In other words, the demand when using HMD+PROJECTOR significantly increased compared to only HMD (except for physical
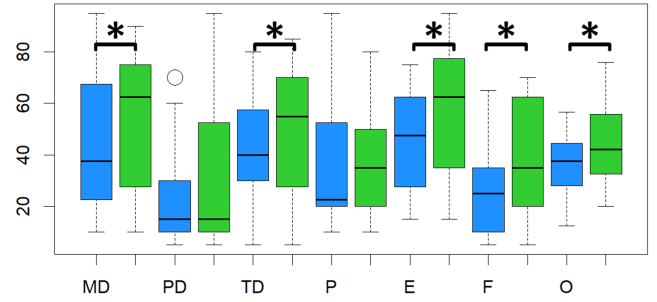
Figure 7: Scores for the NASA TLX questionnaire. From left to right: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (P), Effort (E), Frustration (F) and Overall Demand (O) HMD (blue) and HMD+PROJECTOR (green). Significant differences are marked by a ∗ symbol.

Table 1: Average SSQ results with standard deviation in parenthesis.

| Scale	HMD	HMD+PROJECTOR
Total	31.8	36.5
	(16.0)	(21.9)
Nausea	15.9	18.3
	(15.3)	(18.6)
Oculo-Motor	34.4	39.5
	(19.8)	(21.8)
Disorientation	29.6	36.5
	(20.2)	(30.9)

demand). Also, pariticpant subjectively did not judge their performance significantly different in both conditions.

3.5.4 Simulator Sickness. The overall results for the simulator sickness questionnaire are depicted in Table 1. Wilcoxon signed rank tests indicated no significant differences for the total SSQ score, nor for the subscales nausea, oculo-motor or disorientation.

Please note, that gender-specific differences were indicated for the total score in both HMD ($Z = 2.26, p = 0.025, Cohen's\ d = 1.04$) and HMD+PROJECTOR ($Z = 3.13, p = 0.001, Cohen's\ d = 1.66$) as well as for oculo-motor ($Z = 2.44, p = 0.015, Cohen's\ d = 1.15$) and disorientation ($Z = 2.46, p = 0.015, Cohen's\ d = 1.16$) scores in the HMD+PROJECTOR condition between female and male, but not within each group between HMD+PROJECTOR and HMD, see also Figure 8.

In other words, while gender-specific differences were detected within conditions (as suggested by literature on gender-specific differences in motion sickness [27]), no significant differences were detected between conditions for simulator sickness.

3.5.5 Search Strategies. For condition HMD, 10 out of 24 participants mentioned that they saw if the target was present or not at first glance. Nine participants mentioned that they first focused on the image center and, in case they did not spot the target, searched
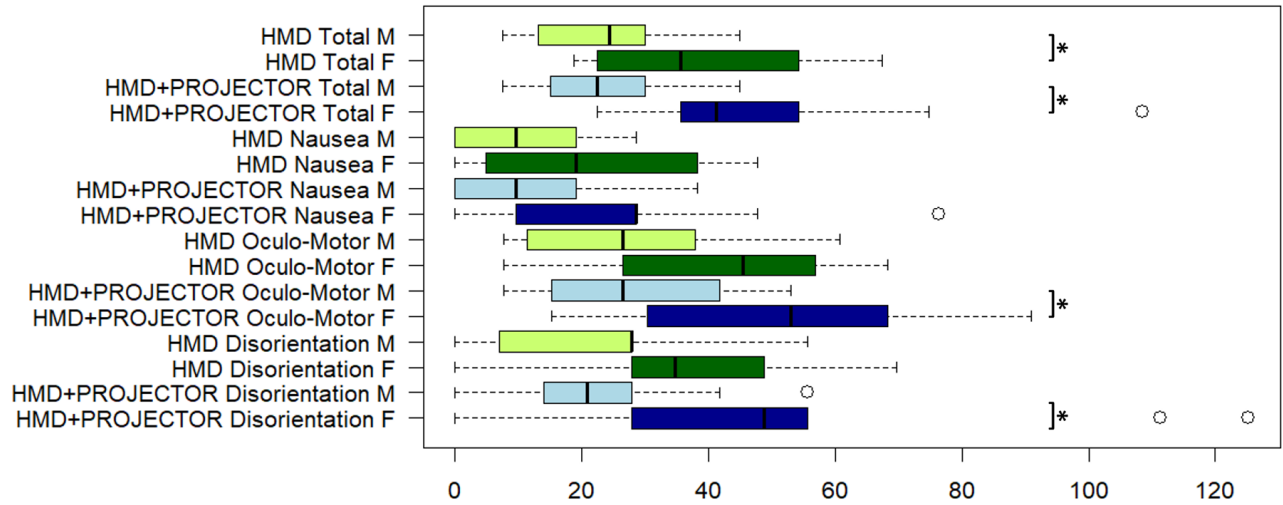
Figure 8: Gender-specific differences of the SSQ questionnaire results. Significant differences are marked by a ∗ symbol. Please note, that no significant difference were found between conditions for male or female participants. F: female. M: male.

the corners and borders in a clockwise order. The remaining participants conducted the search following an S-shape, starting on the top left of the screen, or divided the screen into four quadrants, which they searched subsequently.

For condition HMD+PROJECTOR, six participants mentioned they always focused on the center of the projection area first, and afterwards on the depth layer of the HMD. Four participants mentioned they always focused first on the background layer and thereafter on the projection in the foreground. Three participants mentioned they exclusively focused on the background layer (such that the foreground was always perceived blurry). The remaining 11 participants indicated they regularly switched between focusing first on the foreground or the background.

No significant differences regarding task completion time, errors, or workload were indicated between those groups.

4 DISCUSSION AND IMPLICATIONS

The results indicate that conducting a visual search task across two depth layers results in a significantly higher reaction time with an approximate increase of 50% and significantly higher error rate with an approximately increase of 100% compared to using a single depth layer. In addition, the perceived workload was higher for two depth layers, however simulator sickness revealed no significant difference.

While previous work has shown that there is an effect of depth layer and context switching [15, 24], this experiment quantifies that the reaction time and error rate for joint visual information processing across an OST HMD and a body-proximate display are substantially higher compared to visual information processing at a single depth layer. Factors such as display type (e.g., projector vs. smartphone, different OST-HMD models), display parameters (such as resolution) could influence the absolute values of the results, but we expect the magnitude of the results to not substantially change for the difference in viewing distances evaluated in this experiment.

However, it would be worthwile to better understand how the performance and subjective parameters change as the vergence and accomodation distances converge.

While the general direction of these results was expected, the extent of the negative effects was previously unknown and highlights the need for thoughtful design when distributing information across an OST HMD and a body-proximate display. As we anticipate such solutions to emerge in the near future as mitigation tactics against the user interface limitations inherent in the current generation OST HMDs in practical applications, the extent of the cost of such solutions is valuable to make informed design decisions.

The results also highlight the need for research to both understand the full extent of the consequences of joint OST HMD and body-proximate display systems and to identify successful mitigation techniques and strategies. While some techniques exists for adaptive sharpening of out-of-focus content, e.g., [10, 34], it remains to be seen, if they can have a positive benefits for joint visual information processing tasks.

However, the effect of double vision, when verging the eyes between foreground and background layer, remains an open issue. While this could be theoretically addressed by putting the vergence plane of an OST HMD to 30 cm, in practice this would substantially reduce the effective field of view in which stereo rendered content could be perceived. In the future, vari-focal or multi-focal OST HMDs (e.g., [1, 13, 48]) might potentially mitigate the effects indicated in our study.

One option for current generation OST-HMDs might be to minimize the amount of joint information processing across the two depth layers in safety critical situations using such a multi-display setup. This does not implicate to only use a single depth layer all the time, but to carefully direct the attention of users either to the HMD or to the body-proximate display at a time. However, further studies would need to investigate this in detail.

5 CONCLUSIONS AND FUTURE WORK

Joint OST HMD and body-proximate display systems open up a new solution space. However, there are also likely negative effects, such as reduced visual information integration performance among users. To understand the extent of this cost this paper presents the result of an experiment evaluating human performance in a visual search task across an OST HMD and a body-proximate display at 30 cm.

The results revealed that task completion times significantly increased by approximately 50% and error rates significantly increased by approximately 100% compared to a visual search on a single depth layer.

In addition, perceived workload ratings were higher for joint visual information processing across two depth layers. The results quantify the negative joint visual information integration cost and highlight a design trade-off when exploring joint OST HMD and body-proximate display systems.

In future work, we would like to study the effects of joint information processing acrosss OST HMDs and body-proximate displays in work scenarios, such as following a repair instruction at a physical machine. This would further inform users and manufacturers of current generation HMDs in industrial settings about the expected costs when using HMDs with physical objects within hands reach, specifically as some manufacturers such as Microsoft are pushing OST HMDs as supporting tool for first-line workers [29].

Further, we plan to replicate the study with further OST HMDs such wide field-of-view HMDs (to set the vergence plane at the same distance as the body-proximate display, which would still not resolve the accomodation issue), as well as multi- or varifocal HMDs. Even if the accomodation conflict would be mitigated using those HMDs, research indicates that there still remains a cost for jointly processing the information across an HMD and a further output medium at the same depth layer [24].

REFERENCES

[1] Kaan Akşit, Ward Lopes, Jonghyun Kim, Peter Shirley, and David Luebke. 2017. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 189.

[2] Yuliya Bababekova, Mark Rosenfield, Jennifer E Hue, and Rae R Huang. 2011. Font size and viewing distance of handheld smart phones. *Optometry and Vision Science* 88, 7 (2011), 795–797.

[3] Michael Bach et al. 1996. The Freiburg Visual Acuity Test-automatic measurement of visual acuity. *Optometry and vision science* 73, 1 (1996), 49–53.

[4] Tanja Blascheck, Anastasia Bezerianos, Lonni Besançon, Bongshin Lee, and Petra Isenberg. 2018. Preparing for perceptual studies: Position and orientation of wrist-worn smartwatches for reading tasks.

[5] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Klokmose, and Nicolai Marquardt. 2019. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, New York. ISBN 9781450359702 (In Press)*. ACM.

[6] Rahul Budhiraja, Gun Lee, and Mark Billinghurst. 2013. Using a hhd with a hmd for mobile ar interaction. (2013).

[7] Jessica R Cauchard, Markus Löchtefeld, Pourang Irani, Johannes Schoening, Antonio Krüger, Mike Fraser, and Sriram Subramanian. 2011. Visual separation in mobile multi-display environments. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 451–460.

[8] Xiang'Anthony' Chen, Tovi Grossman, Daniel J Wigdor, and George Fitzmaurice. 2014. Duet: exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 159–168.

[9] Pei-Yu Peggy Chi, Yang Li, and Björn Hartmann. 2016. Enhancing cross-device interaction scripting with interactive illustrations. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 5482–5493.

[10] Trey Cook, Nate Phillips, Kristen Massey, Alexander Plopski, and Christian Sandor. 2018. User Preference for SharpView-Enhanced Virtual Text During Non-Fixated Viewing. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1–400.

[11] David Dobbelstein, Christian Winkler, Gabriel Haas, and Enrico Rukzio. 2017. PocketThumb: a Wearable Dual-Sided Touch Interface for Cursor-based Control of Smart-Eyewear. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 9.

[12] Jakub Dostal, Per Ola Kristensson, and Aaron Quigley. 2013. Subtle gaze-dependent techniques for visualising display changes in multi-display environments. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 137–148.

[13] David Dunn, Cary Tippets, Kent Torell, Petr Kellnhofer, Kaan Akşit, Piotr Didyk, Karol Myszkowski, David Luebke, and Henry Fuchs. 2017. Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. *IEEE transactions on visualization and computer graphics* 23, 4 (2017), 1322–1331.

[14] Stephen R Ellis and Brian M Menges. 1998. Localization of virtual objects in the near visual field. *Human factors* 40, 3 (1998), 415–431.

[15] Joe Gabbard, Divta Gupta Mehra, and J Edward Swan II. 2018. Effects of AR display context switching and focal distance switching on human performance. *IEEE Transactions on Visualization and Computer Graphics* (2018).

[16] Joseph L Gabbard, Gregory M Fitch, and Hyungil Kim. 2014. Behind the Glass: Driver challenges and opportunities for AR automotive applications. *Proc. IEEE* 102, 2 (2014), 124–136.

[17] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. Multifi: Multi fidelity interaction with displays on and around the body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3933–3942.

[18] Jens Grubert and Matthias Kranz. 2017. Headphones: Ad hoc mobile multi-display environments through head tracking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3966–3971.

[19] Jens Grubert and Matthias Kränz. 2017. Towards ad hoc mobile multi-display environments on commodity mobile devices. In *Virtual Reality (VR), 2017 IEEE*. IEEE, 461–462.

[20] Jens Grubert, Matthias Kranz, and Aaron Quigley. 2016. Challenges in mobile multi-device ecosystems. *mUX: The Journal of Mobile User Experience* 5, 1 (2016), 5.

[21] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183. https://dx.doi.org/10.1016/S0166-4115(08)62386-9

[22] Michael E Holmes, Sheree Josephson, and Ryan E Carney. 2012. Visual attention to television programs with a second-screen application. In *Proceedings of the symposium on eye tracking research and applications*. ACM, 397–400.

[23] Steven Houben and Nicolai Marquardt. 2015. Watchconnect: A toolkit for prototyping smartwatch-centric cross-device applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1247–1256.

[24] Anke Huckauf, Mario H Urbina, Jens Grubert, Irina Böckelmann, Fabian Doil, Lutz Schega, Johannes Tümler, and Rüdiger Mecke. 2010. Perceptual issues in optical-see-through displays. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*. ACM, 41–48.

[25] Maria Husmann. 2017. *Investigating Tool Support for Cross-Device Development*. Ph.D. Dissertation. ETH Zurich.

[26] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.

[27] Sibylle Klosterhalfen, Sandra Kellermann, Fang Pan, URSuLA STOcKHORST, Geoffrey Hall, and PAuL ENcK. 2005. Effects of ethnicity and gender on motion sickness susceptibility. *Aviation, space, and environmental medicine* 76, 11 (2005), 1051–1057.

[28] Ernst Kruijff, J Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 3–12.

[29] Microsoft. 2019 (accessed July 7, 2019). New HoloLens 2 gives Microsoft the edge in the next generation of computing. https://news.microsoft.com/innovation-stories/hololens-2/.

[30] Timothy Neate, Matt Jones, and Michael Evans. 2015. Designing attention for multi-screen TV experiences. In *Proceedings of the 2015 British HCI Conference*. ACM, 285–286.

[31] Timothy Neate, Matt Jones, and Michael Evans. 2015. Mediating attention for second screen companion content. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3103–3106.

[32] Michael Nebeling, Elena Teunissen, Maria Husmann, and Moira C Norrie. 2014. XDKinect: development framework for cross-device interaction using kinect. In *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 65–74.

[33] Erwan Normand and Michael J McGuffin. 2018. Enlarging a smartphone with AR to create a handheld VESAD (Virtually Extended Screen-Aligned Display).

In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 123–133.

[34] Kohei Oshima, Kenneth R Moser, Damien Constantine Rompapas, J Edward Swan, Sei Ikeda, Goshiro Yamamoto, Takafumi Taketomi, Christian Sandor, and Hirokazu Kato. 2016. SharpView: Improved clarity of defocussed content on optical see-through head-mounted displays. In *Virtual Reality (VR), 2016 IEEE*. IEEE, 253–254.

[35] Aaron Quigley and Jens Grubert. 2015. Perceptual and social challenges in body proximate display ecosystems. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 1168–1174.

[36] Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. Huddlelamp: Spatially-aware mobile displays for ad-hoc around-the-table collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 45–54.

[37] Umar Rashid, Miguel A Nacenta, and Aaron Quigley. 2012. The cost of display switching: a comparison of mobile, large display and hybrid UI configurations. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, 99–106.

[38] Umar Rashid, Miguel A Nacenta, and Aaron Quigley. 2012. Factors influencing visual attention switch in multi-display user interfaces: A survey. In *Proceedings of the 2012 International Symposium on Pervasive Displays*. ACM, 1.

[39] Nina Rosa, Wolfgang Hürst, Peter Werkhoven, and Remco Veltkamp. 2016. Visuotactile integration for depth perception in augmented reality. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 45–52.

[40] Mario Schreiner, Roman Rädle, Hans-Christian Jetter, and Harald Reiterer. 2015. Connichiwa: a framework for cross-device web applications. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2163–2168.

[41] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. 2015. Gluey: Developing a head-worn display interface to unify the interaction experience in distributed display environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 161–171.

[42] Teddy Seyed, Alaa Azazi, Edwin Chan, Yuxi Wang, and Frank Maurer. 2015. Sod-toolkit: A toolkit for interactively prototyping and developing multi-sensor, multi-device environments. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM, 171–180.

[43] Gurjot Singh, Stephen R Ellis, and J Edward Swan II. 2018. The effect of focal distance, age, and brightness on near-field augmented reality depth matching. *IEEE transactions on visualization and computer graphics* (2018).

[44] Gurjot Singh, J Edward Swan II, J Adam Jones, and Stephen R Ellis. 2010. Depth judgment measures and occluding surfaces in near-field augmented reality. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*. ACM, 149–156.

[45] Maximilian Speicher, Brian D Hall, Ao Yu, Bowen Zhang, Haihua Zhang, Janet Nebeling, and Michael Nebeling. 2018. XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development. *Proceedings of the ACM on Human-Computer Interaction* 2, EICS (2018), 7.

[46] J Edward Swan, Gurjot Singh, and Stephen R Ellis. 2015. Matching and reaching depth judgments with real and augmented reality targets. *IEEE transactions on visualization and computer graphics* 21, 11 (2015), 1289–1298.

[47] Daniel J Weintraub. 1992. *Human factors issues in head-up display design: The book of HUD*. Technical Report. DAYTON UNIV OH RESEARCH INST.

[48] Austin Wilson and Hong Hua. 2018. High-resolution optical see-through vari-focal-plane head-mounted display using freeform Alvarez lenses. In *Digital Optics for Immersive Displays*, Vol. 10676. International Society for Optics and Photonics, 106761J.

[49] Marc D Winterbottom, Robert Patterson, Byron J Pierce, Christine M Covas, and Jennifer Winner. 2007. Depth of focus and visual recognition of imagery presented on simultaneously viewed displays: implications for head-mounted displays. *Human factors* 49, 5 (2007), 907–919.

[50] Dennis Wolf, John J Dudley, and Per Ola Kristensson. 2018. Performance Envelopes of In-Air Direct and Smartwatch Indirect Control for Head-Mounted Augmented Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 347–354.

[51] JS Wolffsohn, GK Edgar, and NA McBrien. 1998. The effect of viewing a car head-up display on ocular accommodation and response times. *Vision in vehicles* 6 (1998), 143–151.