A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration using AR

THOMAS WELLS, Lancaster University, UK
DOMINIC POTTS, Lancaster University, UK
STEVEN HOUBEN, Eindhoven University of Technology, NL

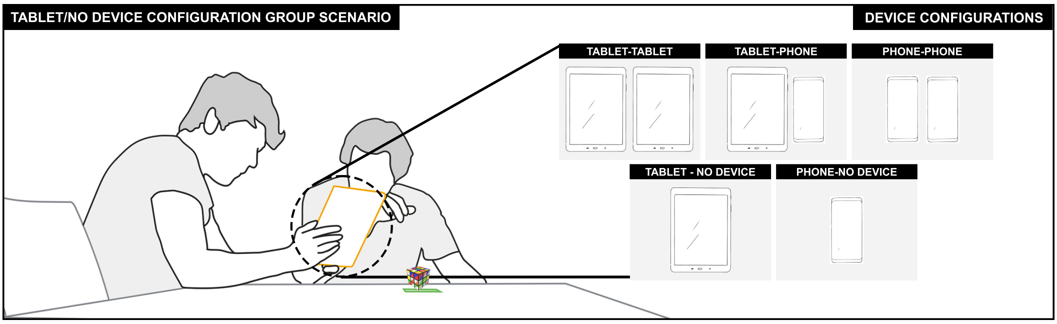

Fig. 1. Using a basic Mobile Augmented Reality interface, we study how five device configurations of varying size and pairing affect group practices through identifying and locating tasks on two virtual models.

The increasing availability of portable handheld mobile Augmented Reality technology is revolutionising the way digital information is embedded into the real world. As this data is embedded, it enables new forms of cross-device collaborative work. However, despite the widespread availability of handheld AR, little is known about the role that device configurations and size play on collaboration. This paper presents a study that examines how completing tasks using a simple mobile AR interface on different device sizes and configurations impacts key factors of collaboration such as collaboration strategy, behaviour, and efficacy. Our results show subtle differences between device size and configurations that have a direct influence on the way people approach tasks and interact with virtual models. We highlight key observations and strategies that people employ across different device sizes and configurations.

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

Additional Key Words and Phrases: Augmented Reality, Mobile Augmented Reality, Mobile Interaction, Collaboration, Co-Located Collaboration

ACM Reference Format:
Thomas Wells, Dominic Potts, and Steven Houben. 2022. A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration using AR. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 200 (September 2022), 23 pages. https://doi.org/10.1145/3546735

Authors' addresses: Thomas Wells, t.wells@lancaster.ac.uk, Lancaster University, Lancaster, UK; Dominic Potts, d.potts2@lancaster.ac.uk, Lancaster University, Lancaster, UK; Steven Houben, s.houben@tue.nl, Eindhoven University of Technology, Eindhoven, NL.

2573-0142/2022/9-ART200 $15.00
https://doi.org/10.1145/3546735

1 INTRODUCTION

Handheld mobile devices have transformed the way Augmented Reality (AR) is used. Due to the widespread availability of AR-capable mobile devices, AR experiences are currently accessible to millions of people worldwide [1] through common consumer applications [38]. Mobile handheld Augmented Reality (or mobile AR) is the most common way people are exposed to augmented content in real world applications [4, 5]. However, despite the ubiquity of mobile AR, there is little empirical or experimental research that investigates how specific mobile AR characteristics and configurations affect collaboration in virtual environments.

One of the central assumptions of mobile AR is that it facilitates and supports collaborative work, as it allows multiple people to interact with virtual content around a shared physical setting [28]. Early systems-oriented work demonstrate that mobile devices can be a powerful tool for collaboration in specific contexts or domains [13, 32, 36, 40]. However, while mobile AR may offer a means to manipulate content augmented onto the real world with fluid interaction, a recent study shows that, without specific features for group work, mobile AR negatively impacts collaboration through increased context switching and a reduction in direct collaboration, as well as influencing how people use their personal devices [48]. There are other assumptions about mobile AR and face-to-face communication around a shared work space - such as its accessibility - but the role of the mobile handheld device itself is not well understood [23]. Currently, no studies or insights look at the role device form factor and configurations play in guiding group collaboration in mobile AR – meaning *we do not understand how the use of various device setups and sizes affects group work.*

Inspired by work on device size in cross-device computing [3, 49], this paper is an initial step in understanding the characteristics of mobile devices. As part of a study aimed at evaluating the impact of screen size and device configuration on AR-mediated co-located collaboration, we designed and conducted an experiment with dyads of participants to examine group strategies and usability metrics as they interacted with a test application. We investigated the effects of five different device configurations spanning two different device sizes (a mobile phone and a tablet) on collaboration with mobile AR. We examined (i) user's ability to complete tasks and the perceived workload, (ii) collaborative behaviours such as focus, communication, and device interaction, as well as (iii) general strategies of collaboration adopted by participants.

Through this study, we aim to understand the implications of device symmetry and sizes on the way task labour is divided in AR for pairs of users in a collaborative tabletop setting. We conducted an experimental study, using a preexisting mobile AR system for model manipulation [48]. Participants were presented with two different models augmented onto the real world (hereby referred to as virtual objects) with different levels of occlusion. Participants were required to complete two types of collaborative tasks on these virtual objects in AR, namely identifying and locating specifics about them. We believe the results provided will help future researchers better understand how screen size affects their subjects, especially when designing and evaluating mobile applications. Our findings show preliminary evidence and insights into the nature of collaboration between device configurations, including device symmetry, size, and quantity, as well as the effect of model occlusion and type of AR activity on collaboration.

2 RELATED WORK

Our study examines how device size and configuration play a role in users' collaboration when using handheld mobile AR. It builds on prior studies that investigate (i) co-located AR collaboration, (ii) the effect of screen size on tasks, and (iii) cross-device systems.

[1] https://www.statista.com/statistics/1098630/global-mobile-augmented-reality-ar-users/

2.1 Co-Located AR Collaboration

Prior research on AR applications have focused on single-user experiences, but the emphasis more recently has been on understanding the implications of collaborative AR, where two or more users experience and interact with the same virtual content. The generally agreed goal of collaborative Augmented Reality is to “augment the face-to-face collaborative experience” [28], yet the collaborative aspect of AR remains under explored [23] and there are plenty of specific areas that require focus [45]. The advent of collaborative AR has introduced new types of informal social interaction in commercial shared AR spaces, most commonly involving the sharing of user experiences. This can be seen in most social media applications [2] [3] [4], which utilise AR stickers and filters to place over the real world. We also see this in persistent AR mobile games, such as Pokemon Go [5] and Minecraft Earth [6], where the content creates a foundation for interaction between multiple users. The use of AR has been shown to reduce cognitive load, task durations, and errors, and enhance learning, facilitating areas such as training, maintenance, and education compared with current methods [12, 20, 27, 35]. According to research on different collaboration methods that include synchronous collaboration among pairs of participants in different mixed reality settings, AR-to-AR settings contributed to increased collaboration, embodiment, and presence [13, 26, 37], and users prefer synchronous, co-located collaboration [9]. However, there is still a lack of understanding as to how differences in device size and configuration could affect collaboration methods and strategies in co-located synchronous AR collaborations.

2.2 The Effect of Screen Size

There are not many studies that deal with the possible effects of screen size, especially where screen-based interactions play a crucial role for example in handheld mobile devices. Studies shows that device size has an effect on tasks. Hancock et al. examined four differing screen sizes under three varying levels of time pressure and found “performance decrement and workload elevation” in small device form factors when under restrictive time pressure [11]. Sanchez et al. report that display size impacts perceived size, which are more accurate when viewed on smaller screens [42, 43]. Maniar et al. examined the effect of non-touch mobile phones’ screen size on video based learning [29]. Using phones with a variety of screen sizes, they found that the smallest screen significantly deteriorated the students’ learning effectiveness. Whilst screens at this time were much smaller, they did not find any significant differences between the phones with the larger displays. Slightly more recently, Kim et al. used three mobile devices (3.5", 5.7", and 9.7") and found that the largest screen led to higher participants enjoyment and found to be significant only between the medium and largest devices. They also report that participants who used the smallest device mentioned they were more likely to use a similar mobile device in the future when compared to the medium device [25]. In a follow up study, Kim and Sundar found that a large screen, when compared to a small screen, “positively influence perceived ease of use – and attitude toward – the device” [24]. In addition to mobiles, we also see comparisons between tablets and tabletops. For example, Zagermann et al. explored how participants performed a sensemaking task with two tablets coupled with a shared horizontal display of varying size [49]. However, it still remains unclear how device size affects how AR mediates co-located collaboration methods and strategies.

[2]https://facebook.com
[3]https://instagram.com
[4]https://snapchat.com
[5]https://pokemongolive.com/en/
[6]https://www.minecraft.net/en-us/article/new-game–minecraft-earth

2.3 Cross-Device Systems

To support cross-device use of interfaces on multiple devices, there have been many frameworks and tools proposed [3]. A particular line of research in cross-device systems is the support of collaborative tasks on interactive devices [1, 15–17, 19, 22, 33, 34]. There are some approaches that combine different sized devices, for example combining tablets with tabletops, and also introducing a tablet into a device ecosystem for use as an overview [2]. However, a study that compares tabletop sizes found that users can efficiently collaborate on tablet-sized devices, without the need for shared displays [49]. Using multiple mobile devices for co-located collaboration is common in Augmented Reality, with a trend showing "that handheld mobile AR has recently become the primary display for AR studies" [5]. However, size differences of devices – such as that between a mobile phone and tablet – as well as the different device configurations these can be used in is yet to be explored for AR, and more specifically for co-located collaboration.

3 MOTIVATION

Previous work illustrates that while AR supports new forms of collaboration [7, 45], it also induces problems and challenges relates to device type, screen form factors and AR activity. In this paper, we extend previous work related to hardware used for accessing augmented reality, along with AR components as well as 3D models that are augmented onto the real world, specifically exploring how these affect co-located AR collaborations. We dissect these challenges in co-located collaboration in AR into five individual research categories that we examine in our study: (i) *device symmetry*, (ii) *device size*, (iii) *device quantity*, (iv) *model occlusion* and (v) *AR activities* (see Figure 2).

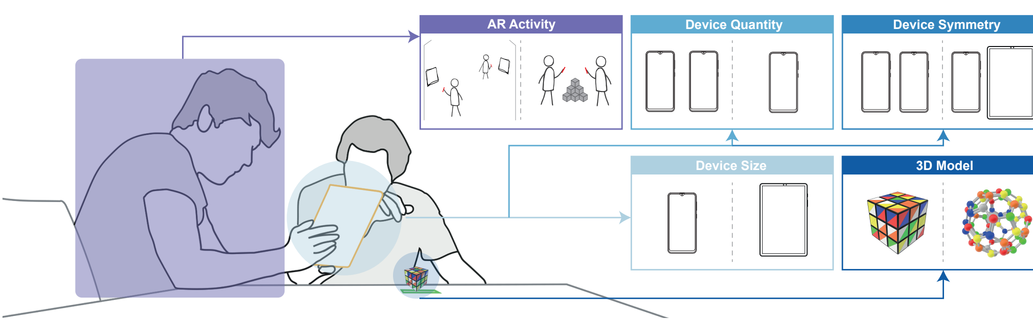

Fig. 2. The five research categories looked into throughout this paper demonstrated in the study set up.

3.1 Device Symmetry (H1)

Device symmetry is a term used frequently in cross-device computing to describe how devices are arranged [30, 41]. In this paper, we refer to device symmetry as when pairs of participants use the same device (i.e. two phones) or different devices (i.e. a tablet and a phone) in a task. When participants need to accomplish a task collaboratively but across separate devices, they may need to take into account differences in asymmetric device configurations. As such, we hypothesise that device asymmetry will impede participant collaboration, with more focus on individual devices in multi-device setups, and increased verbal communication as a workaround.

3.2 Device Size (H2)

Device size is a term commonly used to describe different form factors of device displays. Studies examining the effect of screen size exist (Section 2.2), but the effect of handheld device on collaboration remains poorly explored [3]. Handheld devices can produce a gorilla arm effect [18, 47] depending on their weight, with designers reporting restrictions in mid-air gestures and handheld mobile devices as a result of high consumed endurance and fatigue [8, 14]. Larger devices of course allow more content to be viewed at once, but smaller devices are generally more portable and easier to handle. Our second hypothesis is that larger device sizes will afford more device sharing during collaboration with participants giving more focus to each other's device(s).

3.3 Device Quantity (H3)

Device quantity is simply the number of devices available during collaborative tasks. As the number increases, so does the task of managing the devices in the scenario [10]. In pairs, with a maximum of two devices, the only real consideration is the claiming or sharing of ownership, along with the negotiation of control in multi-device setups. We hypothesise that in single-device configurations, participants with device ownership will interact with the virtual model more, opposed to multi-device setups as participants won't be concerned with breaking shared operational consistency.

3.4 Model Occlusion (H4)

Model occlusion is a problem that occurs in 3D environments and is most commonly seen when a user navigated through them for 3D visualisation. Occlusion of objects is caused by the environment itself, the model, and its geometrical properties [6, 46]. In 3D environments and on 3D models, occlusion effects influence how users perceive concepts such as continuity, proximity and atomic orientation [44]. We hypothesise that a more occluded model will result in less autonomous work, and promote more communication among participants. Increasing the amount of collaboration necessary to complete a task.

3.5 AR Activity (H5)

An AR activity is defined as a task or goal that a participant needs to complete using Augmented Reality. This could be something as simple different ways of experiencing a museum [21] or co-creating persistent structures in AR [9]. While the task may be the same, individual user goals may affect the overall AR activity and the overall collaboration to complete that task. We hypothesise that participants will need to adapt their collaboration strategies for different tasks, but the degree of adaptation will vary depending on the device configuration.

4 STUDY DETAIL

Extending on previous work in this area and within the defined research categories, this paper aims to establish an initial understanding of the role of device size and configuration. We assume that in a world of heterogeneous devices, many real collaborations will involve a variety of device combinations and sizes. Similarly, we can assume that the 3D models being interacted with will vary in terms of complexity and occlusion. Additionally, users' activities can vary even when interaction techniques remain consistent when using handheld AR. Considering these assumptions, what implications do these factors have on the strategy and efficacy of user collaboration in AR?

4.1 Device Ecology & Interaction Techniques

We study 5 different device configurations involving handheld mobile devices and tablets that vary in the number of devices and heterogeneity (see Figure 1). These configurations include: Phone + No Device, Tablet + No Device, Phone + Phone, Tablet + Tablet, and Tablet + Phone. We specifically focus on device configurations *designed for pair work*, as related work demonstrates that pairs are sufficient for evaluating general collaboration styles and strategies [16, 39, 49]. The device configuration, interactions, and tasks are designed specifically for tabletop collaborative settings with AR as a mediating technology. We used touch-based WYSIWYG interaction techniques for rotating and scaling of the virtual model (see Figure 3). Inspired by the CollabAR system used by Wells & Houben [48], we intentionally distribute the model interaction across all devices. Each user can use their device to manipulate the model whilst seeing the changes in real-time. This ensured a more organic development of rules and group strategies. We also utilise marker-based tracking for physical positioning of the virtual model which users could freely manipulate.

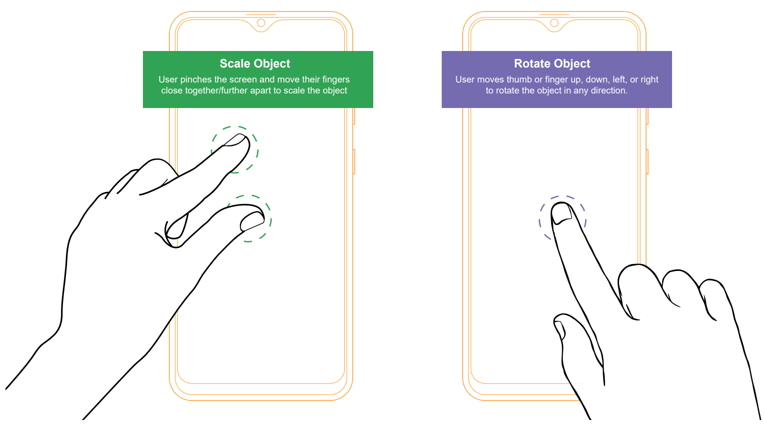

Fig. 3. Basic interactions with the system. Rotation is controlled through swiping on the object and scaling via pinch gestures, as is common in many commercial AR apps.

4.2 AR Tasks & Virtual Models

To evaluate group collaboration, we employ two types of virtual models and two types of task to be completed in every device configuration. The two tasks types were "identify", describe something about the virtual model, and "locate", find an object of interest on the virtual model. These tasks were chosen for their simplicity and application independence. We wanted to ensure the changes to the collaborative components observed were purely due to the device configurations, rather than the complexity of the tasks. The two models, with varying levels of *occlusion*, were used for each device configuration and repeated for each AR task (see Figure 4). Utilising two different models, we can explore the effect of different levels of occlusion within a virtual model on how a device configuration is utilised and how collaboration strategies are formulated. Model 1, a variation on a Rubik's cube, and Model 2, a variation of a fullerene sphere, were designed to be immediately recognisable by participants but with assorted intricate detail, such as randomised colours.

For *identify* tasks, participants would be asked for example *'count how many tiles on the cube that have the colour combination blue and yellow'*. For *locate* tasks, participants would be asked to find something specific such as *'locate the hexagon on the sphere which has the colour combination orange, red, blue, yellow, green, yellow'*. Across both tasks, participants were encouraged to come to a consensus through deliberation before finalising an answer for the task. Task time and task errors were recorded, but participants were informed to complete the tasks within their own time. Once a consensus was reached, participants would proceed with the next task and model.

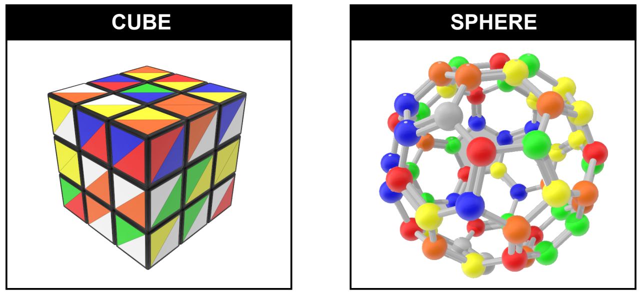

Fig. 4. The two different models used for each task. One with more occlusion (Cube, Left) and one with less occlusion (Sphere, Right).

4.3 Apparatus & Participants

We recruited 20 participants in total (9 identified as male, 11 identified as female) and divided them into pairs. Using a snowball sampling method, we recruited pairs of participants who were already acquainted. There were no special requirements for this study, other than the participants who had no colour vision deficiency, as tasks required the identification of colours. The tasks were abstract and not domain-specific, so no information about the task content was provided prior to the study.

Most participants ages were between 18 and 24, with two participants between the ages of 25 and 34. Whilst not a requirement, all participants had some AR experience prior to this study. Most cited social media filters (such as Snapchat and Instagram) as their use of AR. Other AR applications include Pokémon Go, Google AR, and Minecraft Earth. All participants owned and regularly used at least one handheld device (such as a mobile phone or tablet).

Participants were invited into a circular meeting room setting which allowed for free movement around a table (see Figure 5). Chairs were provided and participants were free to arrange themselves around the table how they pleased. Participants were also informed that they did not have to remain seated. After completing a consent form, participants completed a brief prestudy questionnaire which which gathered basic demographic data, including their prior experience with AR technology.

The device configurations included either a Samsung A70 (mobile device), which has a display size of 6.7", or a Samsung S5 Tab (Tablet), which has a display size of 10.5". The order of device configuration presented to the participants were counterbalanced across each group using the Latin Square method to account for learning effects during the study. In configurations where there was only a single-device, participants were free to decide how ownership and control of the device should work. Each participant was seated around the table, with the experimenter on the opposite end of the room. Pens and NASA-TLX forms were provided to complete after every task. Participants were also free to use the pen and paper for notation when completing the tasks.

After an explanation of the study, participants were given an introduction to the handheld AR system and time to become accustomed to the interaction. After this, they were given a brief sensemaking task which allowed them to familiarise themselves with the models, and also further solidify how to use the system. The introduction was completed using their initial device configuration. Participants were then informed of each task by the experimenter, who also provided the subsequent models by replacing the physical markers. After each task was completed, the participants would complete a post-task NASA-TLX questionnaire to capture the perceived workload for that task using that device configuration, and a likert scale survey regarding how they felt they communicated during this task. After all tasks and configurations were completed, we concluded the study with a semi-structured group interview, probing the different device configurations and the strategies adopted by the participants for the different tasks and models.

Fig. 5. The study setup in a circular meeting room portraying a pair of participants with the Phone + Phone configuration. Virtual models were projected using standard marker tracking.

4.4 Data Collection & Analysis

Each study session was video recorded and the experimenter took observation notes to analyse general **collaborative strategies**, participant **focus**, **communication**, and **group interactions**. After all tasks were completed for a device configuration, the participants were asked to complete a NASA-TLX questionnaire to record their perceived workload for the entire device configuration. In addition, participants would also be asked to complete a likert scale survey to record a self-reflection on how they felt they communicated during this configuration. On completion of all device configurations, the likert scale answers, NASA-TLX questionnaires, and experimenter observations were used to elicit conversation during the semi-structured interview to conclude the study session.

All video footage was analysed by two researchers to define a set of collaboration strategies and, for each participant group, each task was labelled with a predefined strategy. We used and adapted an existing coding framework used by Zagermann et al. [49]. Similarly to this study, we recorded participants interactions with their devices, i.e., how many rotations were performed on the virtual model, as well as their focus and communicative behaviour. The following codes were derived.

0 - Focus - To investigate participant's focus during the tasks, we distinguish between sub codes:
- 0.1 *Focus on Own Device* - where the participant's focus is on their own device.
- 0.2 *Focus on Partner's Device* - where the participant's focus is on their partner's device.
- 0.3 *Focus on Partner* - where the participant's focus is on their partner.
- 0.4 *Focus on Other* - where the participant's focus is on anything else.

1 - Communication - To analyse communication of participants during the study, we coded:
- 1.1 Talking - where the participant is talking.
- 1.2 Silence - where the participant is silent.

We also further coded communicative behaviour as:
- 1.4 Spatial Reference - where the participant makes a verbal reference to the virtual space.
- 1.5 Deictic Gesture - where the participant makes a physical gesture the virtual space.

Participant interaction data was also recorded on the handheld devices during the study. Using this coding scheme, we can determine the level of participant focus and communication for a device configuration at certain intervals in a task or throughout. We can also further describe the nature of collaborative strategies adopted under different device configurations in terms of focus, communication, and interaction. We analyse the NASA-TLX data, task time and error, as well as participant quotations to understand the impact of a device configuration and collaboration strategy on the perceived *workload* and *task efficacy*.

5 RESULTS

The follow section presents the results from the user study with dyads of participants. We categorise these results into findings associated with the efficacy, nature, and strategies of collaboration. We observe a balance between both positive and negative effects that relate to how dyads are affects in their focus, mental and task workload, and general communicative behaviours. We also observe how collaborative strategies are adopted, adjusted, and appropriated to meet the challenges that these effects bring.

5.1 Efficacy of Collaboration

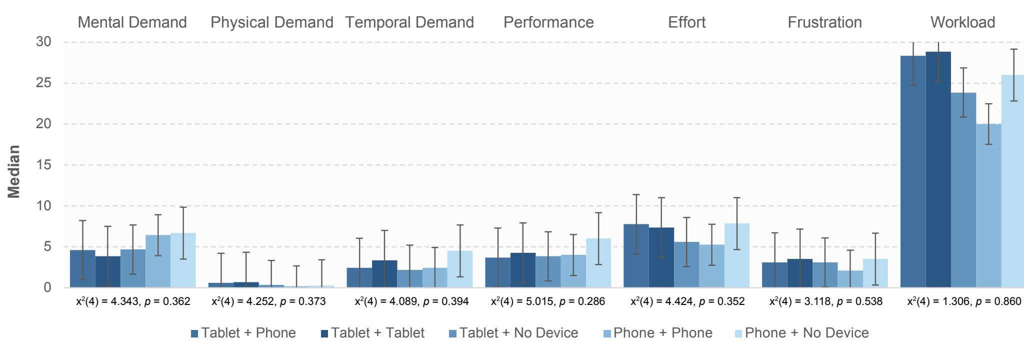

Fig. 6. Median responses for the NASA-TLX questionnaire metrics and overall perceived workload, along with K-Related samples chi square results, where the p-values report the significance between device configurations.

5.1.1 Individual NASA-TLX Metrics. A Friedman test was run on each individual TLX metric and overall Perceived Workload to determine differences exist in reported scores across device configurations. It was found that most metrics were not statistically significant (see Figure 6 for individual test results). We see a trend in the Median values that indicates a higher perceived workload across all metrics on configurations that contain a tablet, including Tablet + Tablet (Mdn = 28.85) and Tablet + Phone (Mdn = 28.35), followed closely by Phone + No Device (Mdn = 26.00). The lowest perceived workload was in Phone + Phone configurations (Mdn = 20.00). Below, we break down each individual metric and note any observed differences across them.

Mental Demand

Configurations that contained tablets had the lowest medians, with the symmetric configuration Tablet + Tablet having the lowest (Mdn = 3.84), followed by Tablet + Phone (Mdn = 4.59) and then Tablet + No Device (Mdn = 4.67). Configurations containing only phones tended toward the highest Median, with Phone + No Device having the highest (Mdn = 6.67) followed closely by Phone + Phone (Mdn = 6.42). We see a large difference between the symmetric Tablet + Tablet configuration and the single Phone + No Device configuration in terms of how mentally demanding they made collaboration.

Physical Demand

We see that the medians for Physical Demand rest mostly the same across all configurations, which lends to the tasks not needing much physical exertion to complete. We do see that configurations containing a tablet tended to slightly higher Physical Demand, such as Tablet + Tablet recorded as the highest (Mdn = .67) and Tablet + Phone recorded as the second highest (Mdn = .58). In group interviews, some participants noted that holding, manoeuvring, and

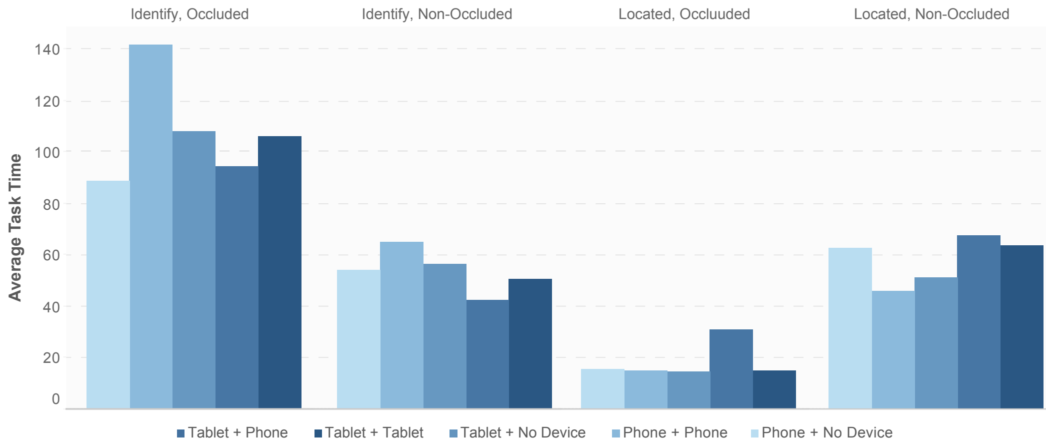

Fig. 7. The mean task time for each device configuration to complete the tasks.

interacting with a larger screen had an impact on their physical demand. This holds true with the median of Phone + Phone configurations (Mdn = .17) reported as the lowest. However, when using the Phone + No Device configuration (Mdn = .25), we observed that participants generally moved around the space more. Primarily to join their partner, but also to look around the model.

Temporal Demand

Participants tended to feel more time pressure in Tablet + Tablet (Mdn = 3.34) and Phone + No Device (Mdn = 4.50) configurations. Other configurations reported roughly the same, where participants did not feel rushed to complete the tasks (see Figure 6).

Performance

Performance metrics with the highest ratings came from the single device configuration Phone + No Device (Mdn = 6.00), and the dual device configuration Tablet + Tablet (Mdn = 4.25). This shows that participants generally felt they performed better when they would either share a device, or had a larger device screen available.

Effort

There is a notable difference between the reported Medians for Tablet + No Device (Mdn = 5.58) and Phone + Phone (Mdn = 5.25) configurations being lower in Effort when compared with the other three configurations. The highest reported Effort was in the Phone + No Device configuration (Mdn = 7.83) with Tablet + Phone (Mdn = 7.75) not far behind. There seems to be no correlation between Effort and specific devices, only that sharing one larger device or having two devices of the same smaller size results in similar Effort.

Frustration

The level of frustration remains mostly the same across all device configurations. There seems to be a noticeable difference between the Phone + Phone (Mdn = 2.09) configuration which has slightly less Median frustration than the other configurations, and the opposite symmetric configuration Tablet + Tablet (Mdn = 3.50) has the highest frustration. This shows that configurations of symmetry has a small difference in frustration between device sizes.

5.1.2 Task Completion Times. We ran a two-way repeated measures ANOVA to determine the effect of different configurations and levels of model occlusion on task completion times. Analysis of the time data showed non-normal distribution (right skewed) when assessed by Shapiro-Wilk's test of normality. As such, we transformed the data using a log transformation and re-assessed using the same test of normality and found it to be normally distributed ($p > .05$) with no outliers. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $X^2(9) = 15.50, p = .085$. There was no statistically significant interaction between model and configuration, $F(4, 36) = .42, p = .796$. However, the main effect of model showed a statistically significant difference in task completion, $F(1, 9) = 96.62, p < .001$, showing that the observation of more occluded models having an increased average task completion time is correct.

Through further observation of Figure 7, we see that by far the shortest task to complete on average was the Locate task on the more occluded object. However, the occluded object also took the longest to complete Identify tasks on. We see that over the Phone + No Device configuration, the Identify tasks generally took the shortest amount of time. Whilst in the dual-device Phone + Phone configuration, it took the most time. Whereas Locate tasks were in fact the shortest time on average in the Phone + Phone configuration, but the most time-consuming in Tablet + Phone configurations. We can infer that, compared to asymmetrical configurations such as Tablet + Phone, symmetrical configurations such as Phone + Phone and Tablet + Tablet took longer on Identify tasks. We also see a trend that **more occluded** models increased the average task completion time in all cases except of the Phone + No Device configuration.

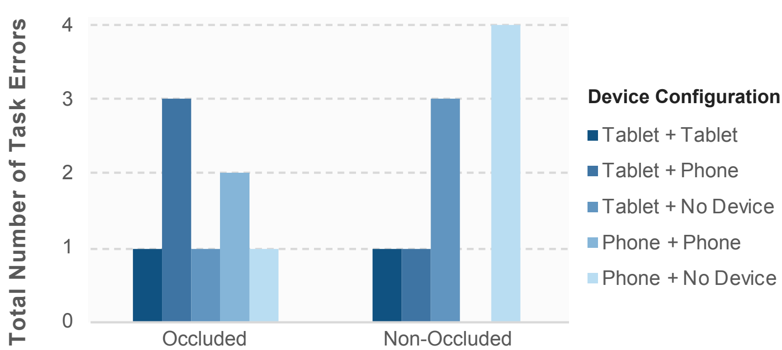

Fig. 8. A total of all errors across all device configurations and models occlusions on the 'Identify' task.

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Device Configuration	-0.071	0.189	0.142	1	0.707	0.931	0.643	1.349
Model Occlusion	-0.142	0.534	0.071	1	0.790	0.868	0.305	2.469
Constant	-1.306	0.661	3.902	1	0.048	0.271		

Table 1. Logistic Regression Predicting Likelihood of Task Error based on Device Configuration and Model Occlusion.

5.1.3 Task Errors. A task error is defined as an incorrect answer given by a group of participants. For task error, we only report the error rate for identify tasks, since locate tasks were not prone to error. We see that both occluded and non-occluded models have a similar number of errors, with the more occluded object having a total of 8 errors, and the less occluded object having a total of 9 errors. A binomial logistic regression was performed to ascertain the effects of device configuration and model occlusion on the likelihood that participants would give an incorrect answer. The logistic regression model was not statistically significant, $X^2(8) = 9.423, p = .899$. Of the two predictor variables, none were statistically significant (as shown in Table 1). Through observation, and if we consider device configuration and task performance as a whole, it was found that while the Phone + No Device configuration had generally shorter task times, they were also the most error prone (see Figure 8). We also see that symmetric configurations Tablet + Tablet and Phone + Phone were the least error prone. Both tablets and phone were similarly prone to errors, but we see a trend in tablets being slightly faster to use.

Fig. 9. A timeline detailing how focus, communication and task completion time varied across groups. Different shades of the same colour represent the two different participants.

5.2 Nature of Collaboration

Following our coding scheme (see Figure 9) outlined in the study detail, we observe how collaborative behaviours change between different device configurations, we analysed participant collaborative behaviours across all configurations, and we categorised our findings into three categories: (i) Focus, (ii) Communication, and (iii) Interaction.

5.2.1 Focus. We ran four Friedman tests to determine if there were differences in Focus across the five different configurations during the tasks. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. It was found that Focus on Own Device was statistically significantly different across different configurations, $X^2(4) = 74.119, p < .001$. Post hoc analysis revealed statistically significant differences between Tablet + No Device ($Mdn = 12.275$) to

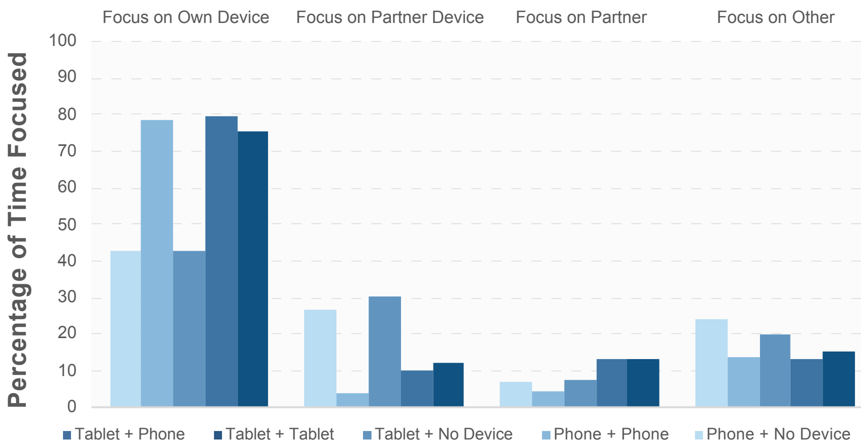

Fig. 10. The average percentage of participant focus during all tasks. 'Focus on other' in this case refers to participants being idle, such as looking around the room.

Phone + Phone ($Mdn = 30.667$)($p < .001$), Tablet + Tablet ($Mdn = 38.576$)($p < .001$), and Tablet + Phone ($Mdn = 37.264$)($p < .001$). Post hoc analysis also revealed statistically significant differences between Phone + No Device ($Mdn = 10.343$) to Phone + Phone ($Mdn = 30.667$)($p < .001$) and Tablet + Phone ($Mdn = 37.264$)($p < .001$). It was also found that Focus on Partner Device was statistically significantly different across different configurations $X^2(4) = 51.206, p < .001$. Post hoc analysis revealed statistically significant differences between Tablet + No Device ($Mdn = 0.000$) to Phone + Phone ($Mdn = 0.000$)($p < .001$), Tablet + Tablet ($Mdn = 0.000$)($p < .001$), and Tablet + Phone ($Mdn = 0.000$)($p < .001$). Post hoc analysis also revealed statistically significant differences between Phone + No Device ($Mdn = 0.000$) to Tablet + Tablet ($Mdn = 0.000$)($p = 0.002$), Phone + Phone ($Mdn = 0.000$)($p = 0.007$), and Tablet + Phone ($Mdn = 0.000$)($p < .001$). No statistically significant differences were found between Focus on Partner or Focus on Other.

Across all device configurations, the participant's primary focus was on their own device without changing their gaze. It was especially prevalent in configurations with two devices, where people spent over 75% of their time looking at their own devices (Tablet + Phone, Tablet + Tablet, Phone + Phone). In these configurations we observed participants looking out their viewports at the object during a discussion even when they coordinated their actions. Sometimes, very briefly, they switch their focus to their partner.

In tasks in which only one device was present, 42% of the time was spent with one participant looking at their own device. Participants who were not in control of the device, however, spent less time concentrating on their partner's device when only a tablet was present, and even less when only a phone was present. When asked about this, participants mentioned that the tablet was "*easier to share*" (P12) because of the larger screen, and that "*phone's are more personal devices*" (P16), so in the real world "*you wouldn't just look over someone's shoulder to view their phone, so it was slightly more awkward*" (P19).

Although there is only a small difference in percentage (around 4%), we see pairs spent more time looking at their own device in configurations that included a phone (Tablet + Phone, Phone + Phone) than in those that included two tablets (Tablet + Tablet). With two phones (Phone + Phone), participants tended to spend most of their time looking at their own devices but very little time

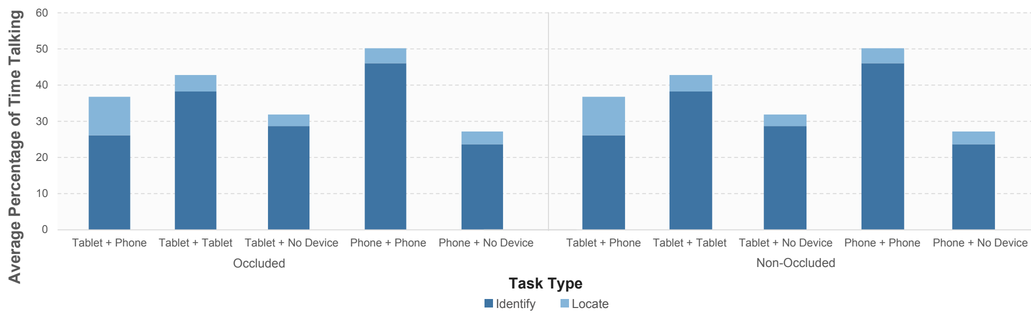

Fig. 11. The average percentage of time participants spent talking in each task.

looking at their partners' devices or their partner. This means they spend most of their time looking around the room, generally waiting for the task their partner is working on to be completed.

The results indicate that in both multi-device and single-phone configurations, participants' attention is drawn more toward their own devices. This is also found in a configuration with a single tablet but to a lesser extent. Due to their focus on their own devices, there is less "face-to-face" interaction. Furthermore, we discovered that all configurations that include a phone seem to draw the user's attention to their own device, lending to the idea that a phone has some innate privacy attached to it, thus making it feel "awkward" (P19) to look over someone's shoulder at their device.

5.2.2 Communication. Based on our observations, we see how communication varies between configurations and tasks (Figure 11). Across the different configurations, we do not observe a large change in the amount of talking. Across all tasks, configurations with two devices had an average talking time of 26%, and configurations with one device had an average talking time of 24%. In the 'Locate' tasks, the configurations with a single device experienced substantially less talk throughout. The participants generally operated from a shared device and adopted a *Independent Work* or a *Asynchronous Work* strategy. Participants felt they "didn't need to discuss" (P6) the model as much as they were "seeing the same thing" (P6). As well as communication across these configurations, we also considered the time participants spent counting out loud, since sometimes this distracted other participants and would result in them losing their count and having to start over. Even with this in mind, the participants discussed very little in their groups, spending an average of 75% of the task time in silence.

We ran a three-way repeated measures ANOVA to determine the effect of different configurations, task, and levels of model occlusion on time spent talking during tasks. Analysis of the time data showed non-normal distribution (right skewed) when assessed by Shapiro-Wilk's test of normality. As such, we transformed the data using a square root transformation and re-assessed them using the same normality test and found them to be normally distributed ($p > .05$) with no outliers. There was no statistically significant difference between configuration, task, and level of occlusion, $F(4, 76) = 1.72, p = .152$.

However, there was a statistically significant simple two-way interaction between task and occlusion, $F(1, 19) = 191.839, p < .001$. Pariwise comparisons in this two-way interaction show that there was statistically significant difference on time talking on the occluded model between the Identify task ($Mean = 5.190$) and Locate task ($Mean = 3.387$)($p < .001$). Pairwise comparisons in this two-way interaction also show that there was statistically significant difference on time talking on the Identify task between the Occluded ($Mean = 5.190$) and Non-Occluded Models

$(Mean = 3.387)(p < .001)$, and on the Locate task between Occluded $(Mean = 1.958)$ and Non-Occluded Models $(Mean = 3.600)(p < .001)$

There was also a statistically significant difference between in time spent talking between configuration and task, $F(4, 76) = 3.788, p = .007$. Pairwise comparisons show statistically significant differences in Phone + No Device to tasks Identify $(Mean = 4.289)(p = .002)$ and Locate $(Mean = 2.779)(p = .002)$, Phone + Phone to tasks Identify and Locate $(p < .001)$, Tablet + No Device and tasks Identify and Locate $(p < .001)$, and Tablet + Tablet to tasks Identify and Locate $(p = .003)$. Pairwise comparisons on time spent talking during Locate tasks found statistically significant difference between Phone + No Device $(Mean = 2.396)$ to Tablet + Phone $(Mean = 3.603)(p = .030)$, and Tablet + No Device $(Mean = 2.311)$ to Tablet + Phone $(Mean = 3.603)(p = .015)$.

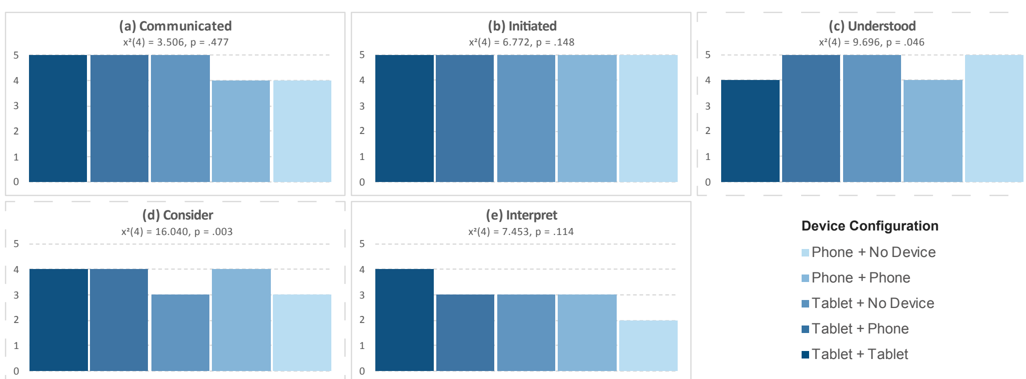

Fig. 12. Median scores on a 5-point Likert scale of participant's individual self-reflection on their communication. A dashed border is present to show significant differences in 'Consider' and 'Understood' questions.

When asked in a Likert-scale questionnaire how participants felt they communicated, across all tasks each participant generally Strongly Agreed that they communicated well with their partner and found it easy to initiated conversation (Figure 12a and b) on tablets, and mostly agreed that they communicated well on phone configurations (Phone + Phone and Phone + No Device). When asked if they felt their partner understood them at all times, we see that there is a trend in dual-device configurations (Tablet + Tablet and Phone + Phone) that they are slightly less agreeable.

A Friedman test was conducted to determine whether agreement in a self-reflection of communication differ between the different device configurations. The results show non-significant differences for how they felt they communicated, how easy they felt it was to initiate conversation, and how carefully they had to interpret their partner's meanings, so therefore retain the null hypothesis. However, we did find significant difference when asked if the participant's needed to carefully consider how to portray their thoughts (Figure 12d) $(x^2(4) = 16.040, p = 0.003)$ and how well they felt their partner understood them $(x^2(4) = 9.696, p = 0.046)$. The results indicate that on configurations that only contain a tablet (Tablet + Tablet and Tablet + No Device), participants had to more carefully consider what they would say to their partner to portray their thoughts. However, separately, on symmetric configurations Tablet + Tablet and Phone + Phone, participants felt less strongly that their partners understood what they were trying to communicate.

Participants used deictic gestures most often to point at their devices. This happened often when participants were working on Identify tasks, as they pointed at the model and counted individual

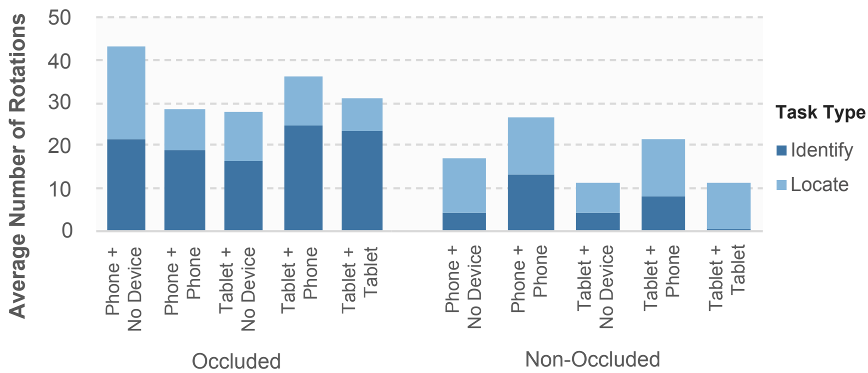

Fig. 13. The average amount of rotations per task on occluded and non-occluded models.

objects. Participants displayed this behaviour primarily when both were focusing on a single device and when referring to a specific area on the virtual model. There were also instances when this happened naturally even when participants were not sharing a view with one another.

5.2.3 Interaction. The occluded model generally required more rotations to complete the task, as opposed to the less occluded model. According to our observation, on the less occluded model, using the symmetric configuration with tablets (Tablet + Tablet), there was a low amount of average rotations. Whereas we saw high amounts of average rotations on the more occluded object in the same configuration. We can see from our video observations that participants would typically scale the non-occluded model larger on most configurations and would be able to see through the object and identify the areas of interest more easily. As opposed to the more occluded object, where participants would be required to rotate the model to cover all areas for both types of tasks.

In general, configurations that included a phone had more interactions than configurations with a tablet. However, one interesting observation was that participants often preferred to use the tablet over the phone in the Tablet + Phone configuration when rotating more occluded models. The participants commented that they preferred using the tablet because they could "see more" (P8) and you could "see rotations better" (P12) during the Tablet + Phone configuration. Regarding other configurations, participants commented that the phone was easier to manipulate the object with due to its size, mentioning “being able to hold it in one hand and rotate it using your thumb” (P1) and that the participants were also “used to using it” in every day life (P14).

5.3 Strategies of Collaboration

A post hoc analysis of participant video data and an examination of collaboration metrics across all tasks was conducted to observe the effects of different device configurations on participant’s general collaboration strategies. Four distinct collaboration strategies emerged:

1 – Independent Work

A task was completed by one participant within the pair *entirely independently*, with their primary **focus on their own device**, little to no **communication**, sharing of device, or negotiation of control in multi-device configurations.

2 – Asynchronous Work
A task was completed independently by each participant *asynchronously*, with primary **focus on their own device**, by sharing device ownership in single-device configurations, or by negotiating control through turn turn-taking in multi-device configurations.

3 – Synchronous Work
Participants completed the task *synchronously* either by sharing a device view in single-device configurations – demonstrated by higher **focus on partner's device** – or **more communicating** with their partner to negotiate control of the virtual object in a multi-device configuration.

4 – Divided Work
A task is explicitly divided into smaller components, often spatially, with participants taking ownership of a component and then working on this simultaneously with more **focus on own devices**. For example, spatially dividing a virtual model into separate areas for each participant to work on for either single or multi-device configurations.

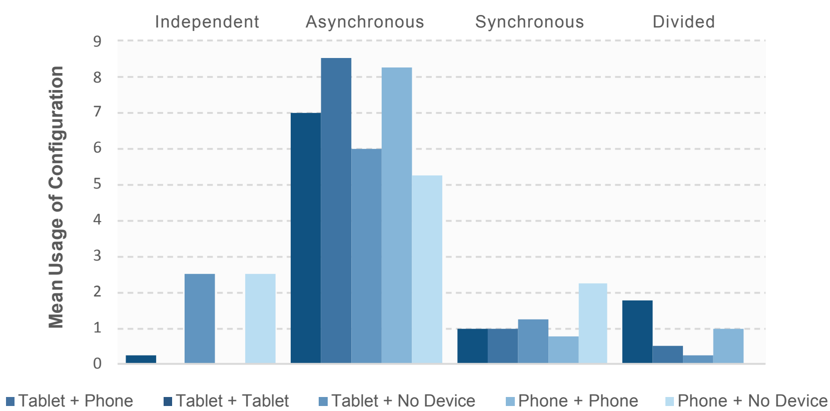

Fig. 14. The average amount of collaboration strategies adopted over each device configuration.

5.3.1 Frequency of strategies across device configurations. Depending on the device configuration, task type, and model type these strategies were applied in varying frequencies (Figure 14).

We see throughout our results that *Synchronous Work* was overwhelmingly the most popular strategy amongst participants across all device configurations, task, and model occlusions. We saw a reduction in the *Synchronous Work* strategy on the Tablet + Phone configuration specifically when participants were working on *Identify* tasks on models that had *more occlusion*. We also see a reduction in this strategy when single-device configurations are in use (Tablet or Phone + No Device). While *Synchronous Work* was the most popular overall for single-device configurations, *Independent Work* and *Asynchronous Work* were adopted more than multi-device configurations and were mostly equal in popularity for both Phone + No Device and Tablet + No Device.

While *Divided Work* was the least popular, there were cases where this strategy increased in popularity. For instance, participants would more often use this approach in multi-device configuration and especially in cases where the devices were symmetrical such as Phone + Phone and Tablet + Tablet. Furthermore, *Divided Work* was used more when working with **more occluded** models, in which case participants would always be positioned face-to-face [31].

6 DISCUSSION

In this research, we have explored the impact of device configuration on collaboration efficacy, behaviour over time, and strategies formed. Our findings, based on five separate research categories, provide insights into the nature of collaboration between different device configurations.

We first hypothesised (H1) that having asymmetric devices would yield more focus on individual devices, and an increase in communication as a workaround. We found that participants tended to focus on one device during single-device configurations (one phone or one tablet) and share. However, the results show that there is still a slightly higher 'Focus on Other' during tasks, which indicate that in some groups, there is still a lot of time spent on individual work, with one participant doing their own thing not in the AR space, which is in support of the hypothesis. This is further backed up by 'Independent Work' being higher on single-device configurations in general. We also see that single-device configurations had statistically significant different to asymmetric multi-device configurations in communication on Locate tasks. However, communication was generally lower on single-device configurations.

Our second hypothesis (H2) was that a larger device size would afford more sharing during collaboration, giving more focus on own and partner devices. This is supported by a trend of somewhat more 'Focus on Partner' in tablet configurations, but no statistically significant differences were detected. While there was no statistically significant difference, we did see that configurations containing a tablet have a higher 'Focus on Partner Device' than symmetric phone configurations. However, there is insufficient statistical difference to conclude that device size facilitates more sharing during collaboration, therefore we reject the second hypothesis.

For our third hypothesis (H3) explains that single-device configurations would lend to increased interaction. Through results, we can observe that the single-device configuration containing a phone had more interactions than any other, while the single-device configurations containing a tablet had one of the lowest across both levels of model occlusion, meaning we reject this hypothesis. This can be related to the additional effort that is necessary to hold a tablet up, while also interacting with the touch screen.

In our fourth hypothesis (H4), we wrote that more occlusion would mean more communication between participants. We observe that with more occluded models, participants would generally position themselves face-to-face. We also saw an increase in Divided Work adoption and less independent work. However, we do not see a large change in the amount of talking when the occlusion changes across the different device configurations.

We hypothesised in our fifth and final hypothesis (H5) that collaborative strategies would need to be adapted depending on the type of task the participants were required to perform. We observe that the most common way participants would collaborate is through a Synchronous Work approach. Once adopted, this strategy would be carried throughout each configuration. The exception being in single-device configurations, in which we see an increase in the Asynchronous Work strategy. Therefore, we reject this hypothesis.

6.1 Task Performance

Our results indicate a trend that task performance was better for symmetric configurations and single-device configurations, though single-device configurations were more prone to error. Observations indicate a reason for this could be due to the higher adoption of *Independent Work* approached for the single-device configurations which — in general — required less communication and negotiation of control. However, it has the disadvantage of tasks not being validated by another user. The increase in task performance in symmetric configurations could be the result of an equal perception of control from participants where *Divided Work* could be more readily applied to decrease the task space, and workload, of each participant, which in turn reduces the likelihood of errors within the task.

Further, the increased perceived workload and frustration seen in all phone-based configurations cannot be attributed to any noteworthy differences in collaboration strategy when compared to tablet-based configurations. However, we can speculate that the form factor and decreased screen 'real estate' of the phone may incur extra physical or cognitive steps for the user. Additionally, the much lower perceived workload of single-device configuration Tablet + No Device might be the result of better affordance of the tablet to the popular *Synchronous Work* collaboration strategy with the increased screen 'real estate' reducing the strain of viewing and perceiving the virtual model simultaneously. Interestingly, configurations that were rated as requiring more effort and more demanding, there was a small increase in participants' perceiving their performance as better.

6.2 Impact of Device on Collaboration

Phones tended to facilitate slightly more focus on their own device, as opposed to tablets which afforded more screen sharing and shoulder-surfing, however these differences are relatively minor. We also see differences in interactions across device configurations. In asymmetric configurations participants, depending on the device ownership and model's occlusion level, would adopt loose roles for completing the task and also manipulating the model which is reflected in the average number of rotations in the Tablet + Phone configuration. For example, tablets were used more often to control more occluded models, whereas phones were used to control less occluded models.

6.3 Collaborative Strategies

Using the *Synchronous Work* strategy, negotiation of control was typically an ongoing and fluid process between participants based on verbal communication, implicit markers such as movement of the model or the other participant, and deictic gestures. This was also similar for *Divided Work*, opposed to *Independent Work* and *Asynchronous Work* where control was often negotiated by device ownership. For both *Synchronous Work* and *Divided Work*, participants occasionally adopted supplementary tools to sketch, annotate, or spatially point on the device screen or in physical space all in reference to the virtual model. The prevalence and popularity of *Synchronous Work* over a similar strategy such as *Divided Work* could be due to the added overhead of democratising the labour of a task. Due to the simplicity or low workload of the tasks, the ad-hoc nature of *Synchronous Work* may have been preferred as a 'quick-and-easy' approach as opposed to something more structured such as *Divided Work*. Despite *Synchronous Work* being the most popular collaborative strategy, participants would encounter issues not observed in other strategies. For example, in single-device configurations, participants sharing a view would be less strict about negotiating model control and would often cause unintended interactions.

For *Independent Work*, participants without device ownership were often comfortable letting the other participant complete the task and trusting the outcome while they were mostly idle, with instances where they would even distract the other participant from completing the task. For *Asynchronous Work*, control of the model was mostly negotiated via device ownership, with the occasional instance of the device handler holding the device for the other participant and inviting them to complete the task and interact. Regarding the type of task, **identify** tasks involved more *Asynchronous Work* and **locate** tasks had more *Independent Work* strategies. This could be due to the lower cognitive effort required for the locate tasks, meaning that *Independent Work* strategies are more desirable. For *Divided Work*, we can postulate that participants were more comfortable using this strategy in symmetric device configurations as there was more perceived equity in terms of the tools available to each participant, despite the interactions being the same in asymmetric device configurations. We can also infer that participants were more encouraged to adopt this strategy if they were not able to perceive most of the information, i.e., when working with a more occluded model.

In summary, we observed and identified distinct collaborative strategies used by participants using handheld AR, and they adopted them at varying rates depending on device configuration. The effect of task type and model occlusion on collaboration strategies and user mobility was also observed. Generally, the results of this study do not make a strong claim that different categories influence co-located collaboration. However, the preliminary results provided insight and evidence about the nature of collaboration between devices of different configurations, including their symmetry, size, and quantity. In addition, there is evidence that model occlusion may affect collaboration, though further research is required.

7 LIMITATIONS & FUTURE WORK

Our work attempts to provide an initial investigation into the influence of different device ecologies and configurations on collaboration in co-located handheld AR. However, there are some limitations with our current work that show potential opportunities for future work to be built on. Firstly, our work primarily focused on tasks and interactions with a 3D virtual model, but AR can be used in different ways to support activities in the physical space, providing additional digital information and content. Naturally, it would be interesting to see how collaboration strategies, behaviour, and effectiveness are influenced when activities include physical and virtual interactions. The tasks and models used in the study were abstract and limited in scope, but provide initial generalisable results for more application and context-specific work to build on.

Furthermore, the tasks were intentionally designed to be possible independently or collaboratively, to keep the participant strategies open, sometimes resulting in less collaborative approaches from the participants. Future work could focus on tasks that require more than one participant to complete, ensuring that in every instance, some manner of collaboration occurs. Regarding the analysis of collaboration, future work could better cross-reference strategy, behaviour, and task efficacy together to increase understanding around which device configurations and strategies work in different tasks and application contexts. Finally, the analysis of the participant interaction could be analysed along with the mobility, spatial formation, and movement of the participant in similar handheld AR tasks to understand how collaboration strategies relate to on-screen interaction and group movement during a task.

8 CONCLUSION

Augmented Reality has the potential to become an instrumental tool for collaboration. While many new technologies and devices are being introduced, currently the dominant way to use and collaborate in AR is through mobile devices. Understanding the precise role that device size and configurations play in mobile handheld AR systems for collaboration is essential for developing new design approaches that enable and facilitate collaboration. In our study, we examine how mobile AR affects (i) efficacy of collaborative work, (ii) changing collaboration behaviour, and (iii) collaboration strategies. Our findings show a nuanced balanced between positive and negative effects related to how participants were affected in their focus, mental/task load, and communication – but also how collaborative settings were adjusted, appropriated and adapted to adjust for such challenges.

ACKNOWLEDGMENTS

We would like to thank our participants for their time, and also our anonymous reviewers for their detailed comments and feedback on this manuscript.

REFERENCES

[1] Christopher Andrews, Alex Endert, and Chris North. 2010. Space to Think: Large High-Resolution Displays for Sensemaking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 55–64. https://doi.org/10.1145/1753326.1753336

[2] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. 2018. Investigating the Role of an Overview Device in Multi-Device Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3173874

[3] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted 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, Glasgow Scotland Uk, 1–28. https://doi.org/10.1145/3290605.3300792

[4] Alan B. Craig. 2013. Chapter 7 - Mobile Augmented Reality. In *Understanding Augmented Reality*, Alan B. Craig (Ed.). Morgan Kaufmann, Boston, 209–220. https://doi.org/10.1016/B978-0-240-82408-6.00007-2

[5] Arindam Dey, Mark Billinghurst, Robert W. Lindeman, and J. Edward Swan. 2018. A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI* 5 (2018).

[6] Niklas Elmqvist and Philippas Tsigas. 2008. A Taxonomy of 3D Occlusion Management for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 5 (Sept. 2008), 1095–1109. https://doi.org/10.1109/TVCG.2008.59

[7] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting Collaboration through Mixed Reality: The Evolution of Groupware. *International Journal of Human-Computer Studies* 131 (Nov. 2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011

[8] Eg Su Goh, Mohd Shahrizal Sunar, and Ajune Wanis Ismail. 2019. 3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review. *IEEE Access* 7 (2019), 40581–40601. https://doi.org/10.1109/ACCESS.2019.2906394

[9] Anhong Guo, Ilter Canberk, Hannah Murphy, Andrés Monroy-Hernández, and Rajan Vaish. 2019. Blocks: Collaborative and Persistent Augmented Reality Experiences. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (Sept. 2019), 83:1–83:24. https://doi.org/10.1145/3351241

[10] Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: Enabling and Understanding Cross-Device Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2773–2782. https://doi.org/10.1145/2556288.2557170

[11] P.A. Hancock, B.D. Sawyer, and S. Stafford. 2015. The Effects of Display Size on Performance. *Ergonomics* 58, 3 (March 2015), 337–354. https://doi.org/10.1080/00140139.2014.973914

[12] Steven Henderson and Steven Feiner. 2011. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE transactions on visualization and computer graphics* 17, 10 (Oct. 2011), 1355–1368. https://doi.org/10.1109/TVCG.2010.245

[13] Anders Henrysson, Mark Billinghurst, and Mark Ollila. 2005. Face to Face Collaborative AR on Mobile Phones. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '05)*. IEEE Computer Society, USA, 80–89. https://doi.org/10.1109/ISMAR.2005.32

[14] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 1063–1072. https://doi.org/10.1145/2556288.2557130

[15] Leila Homaeian, Nippun Goyal, James R. Wallace, and Stacey D. Scott. 2018. Group vs Individual: Impact of TOUCH and TILT Cross-Device Interactions on Mixed-Focus Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173647

[16] Petra Isenberg, Danyel Fisher, Sharoda A. Paul, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. 2012. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (May 2012), 689–702. https://doi.org/10.1109/TVCG.2011.287

[17] Mikkel R. Jakobsen and Kasper HornbÆk. 2014. Up Close and Personal: Collaborative Work on a High-Resolution Multitouch Wall Display. *ACM Transactions on Computer-Human Interaction* 21, 2 (Feb. 2014), 11:1–11:34. https://doi.org/10.1145/2576099

[18] Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling Cumulative Arm Fatigue in Mid-Air Interaction Based on Perceived Exertion and Kinetics of Arm Motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 3328–3339. https://doi.org/10.1145/3025453.3025523

[19] Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer, and Natasa Milic-Frayling. 2011. Materializing the Query with Facet-Streams: A Hybrid Surface for Collaborative Search on Tabletops. In *Proceedings of the SIGCHI*

Conference on Human Factors in Computing Systems (CHI '11). Association for Computing Machinery, New York, NY, USA, 3013–3022. https://doi.org/10.1145/1978942.1979390

[20] Jérome Jetter, Jörgen Eimecke, and Alexandra Rese. 2018. Augmented Reality Tools for Industrial Applications: What Are Potential Key Performance Indicators and Who Benefits? *Computers in Human Behavior* 87 (Oct. 2018), 18–33. https://doi.org/10.1016/j.chb.2018.04.054

[21] Jens Keil, Laia Pujol, Maria Roussou, Timo Engelke, Michael Schmitt, Ulrich Bockholt, and Stamatia Eleftheratou. 2013. A Digital Look at Physical Museum Exhibits: Designing Personalized Stories with Handheld Augmented Reality in Museums. In *2013 Digital Heritage International Congress (DigitalHeritage)*, Vol. 2. 685–688. https://doi.org/10.1109/DigitalHeritage.2013.6744836

[22] Imran Abbas Khawaja, Adnan Abid, Muhammad Shoaib Farooq, Adnan Shahzada, Uzma Farooq, and Kamran Abid. 2020. Ad-Hoc Collaboration Space for Distributed Cross Device Mobile Application Development. *IEEE Access* 8 (2020), 62800–62814. https://doi.org/10.1109/ACCESS.2020.2980319

[23] Kangsoo Kim, Mark Billinghurst, Gerd Bruder, Henry Been-Lirn Duh, and Gregory F. Welch. 2018. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (Nov. 2018), 2947–2962. https://doi.org/10.1109/TVCG.2018.2868591

[24] Ki Joon Kim and S. Shyam Sundar. 2014. Does Screen Size Matter for Smartphones? Utilitarian and Hedonic Effects of Screen Size on Smartphone Adoption. *Cyberpsychology, Behavior, and Social Networking* 17, 7 (July 2014), 466–473. https://doi.org/10.1089/cyber.2013.0492

[25] Ki Joon Kim, S. Shyam Sundar, and Eunil Park. 2011. The Effects of Screen-Size and Communication Modality on Psychology of Mobile Device Users. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 1207–1212. https://doi.org/10.1145/1979742.1979749

[26] K. Kiyokawa, M. Billinghurst, S.E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. 2002. Communication Behaviors of Co-Located Users in Collaborative AR Interfaces. In *Proceedings. International Symposium on Mixed and Augmented Reality*. 139–148. https://doi.org/10.1109/ISMAR.2002.1115083

[27] Wenkai Li, A. Y. C. Nee, and S. K. Ong. 2017. A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. *Multimodal Technologies and Interaction* 1, 3 (Sept. 2017), 17. https://doi.org/10.3390/mti1030017

[28] Stephan Lukosch, Mark Billinghurst, Leila Alem, and Kiyoshi Kiyokawa. 2015. Collaboration in Augmented Reality. *Computer Supported Cooperative Work (CSCW)* 24, 6 (Dec. 2015), 515–525. https://doi.org/10.1007/s10606-015-9239-0

[29] Nipan Maniar, Emily Bennett, Steve Hand, and George Allan. 2008. The Effect of Mobile Phone Screen Size on Video Based Learning. *Journal of Software* 3, 4 (April 2008), 51–61. https://doi.org/10.4304/jsw.3.4.51-61

[30] Nicolai Marquardt, Frederik Brudy, Can Liu, Ben Bengler, and Christian Holz. 2018. SurfaceConstellations: A Modular Hardware Platform for Ad-Hoc Reconfigurable Cross-Device Workspaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–14. https://doi.org/10.1145/3173574.3173928

[31] Paul Marshall, Yvonne Rogers, and Nadia Pantidi. 2011. Using F-formations to Analyse Spatial Patterns of Interaction in Physical Environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work - CSCW '11*. ACM Press, Hangzhou, China, 445. https://doi.org/10.1145/1958824.1958893

[32] D. Mogilev, K. Kiyokawa, M. Billinghurst, and J. Pair. 2002. AR Pad: An Interface for Face-to-Face AR Collaboration. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems (CHI EA '02)*. Association for Computing Machinery, New York, NY, USA, 654–655. https://doi.org/10.1145/506443.506530

[33] Meredith Ringel Morris, Danyel Fisher, and Daniel Wigdor. 2010. Search on Surfaces: Exploring the Potential of Interactive Tabletops for Collaborative Search Tasks. *Information Processing & Management* 46, 6 (Nov. 2010), 703–717. https://doi.org/10.1016/j.ipm.2009.10.004

[34] Meredith Ringel Morris, Jarrod Lombardo, and Daniel Wigdor. 2010. WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 401–410. https://doi.org/10.1145/1718918.1718987

[35] U. Neumann and A. Majoros. 1998. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*. 4–11. https://doi.org/10.1109/VRAIS.1998.658416

[36] Susanna Nilsson, Bjorn Johansson, and Arne Jonsson. 2009. Using AR to Support Cross-Organisational Collaboration in Dynamic Tasks. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. 3–12. https://doi.org/10.1109/ISMAR.2009.5336522

[37] Ye Pan, David Sinclair, and Kenny Mitchell. 2018. Empowerment and Embodiment for Collaborative Mixed Reality Systems. *Computer Animation and Virtual Worlds* 29, 3-4 (2018), e1838. https://doi.org/10.1002/cav.1838

[38] Tim Perdue. 2020. All You Wanted to Know About Augmented Reality. https://www.lifewire.com/applications-of-augmented-reality-2495561.

[39] Thomas Plank, Hans-Christian Jetter, Roman Rädle, Clemens N. Klokmose, Thomas Luger, and Harald Reiterer. 2017. Is Two Enough? ! Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 4548–4560.

[40] Erik Prytz, Susanna Nilsson, and Arne Jönsson. 2010. The Importance of Eye-Contact for Collaboration in AR Systems. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. 119–126. https://doi.org/10.1109/ISMAR.2010.5643559

[41] Roman Rädle, Hans-Christian Jetter, Mario Schreiner, Zhihao Lu, Harald Reiterer, and Yvonne Rogers. 2015. Spatially-Aware or Spatially-agnostic? Elicitation and Evaluation of User-Defined Cross-Device Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3913–3922.

[42] C.A. Sanchez, T. Read, and A. Crawford. 2021. Smartphone Display Size Can Influence Perceptual Judgments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 65, 1 (Sept. 2021), 963–967. https://doi.org/10.1177/1071181321651334

[43] Christopher A. Sanchez, Tyler Read, and Amanda Crawford. 2021. Smartphone Display Size Can Cause Distortions in Perceptual Estimates of Size. *Applied Ergonomics* 97 (Nov. 2021), 103524. https://doi.org/10.1016/j.apergo.2021.103524

[44] Kim Sauvé, Dominic Potts, Jason Alexander, and Steven Houben. 2020. A Change of Perspective: How User Orientation Influences the Perception of Physicalizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3313831.3376312

[45] Mickael Sereno, Xiyao Wang, Lonni Besancon, Michael J Mcguffin, and Tobias Isenberg. 2020. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3032761

[46] Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, and Hans Gellersen. 2020. Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. https://doi.org/10.1145/3313831.3376438

[47] Computer Dictionary of Information Technology. 2022. Gorilla Arm. https://www.computer-dictionary-online.org/definitions-g/gorilla-arm.html.

[48] Thomas Wells and Steven Houben. 2020. CollabAR – Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[49] Johannes Zagermann, Ulrike Pfeil, Roman Rädle, Hans-Christian Jetter, Clemens Klokmose, and Harald Reiterer. 2016. When Tablets Meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 5470–5481. https://doi.org/10.1145/2858036.2858224

Received February 2022; revised May 2022; accepted June 2022