When Tablets meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets

Johannes Zagermann¹, Ulrike Pfeil¹, Roman Rädle¹, Hans-Christian Jetter², Clemens Klokmose³, and Harald Reiterer¹

¹HCI Group, University of Konstanz, {johannes.zagermann,ulrike.pfeil,roman.raedle,harald.reiterer}@uni.kn
²University of Applied Sciences Upper Austria, hans-christian.jetter@fh-hagenberg.at
³Department of Computer Science, Aarhus University, clemens@cs.au.dk

ABSTRACT
Cross-device collaboration with tablets is an increasingly popular topic in HCI. Previous work has shown that tablet-only collaboration can be improved by an additional shared workspace on an interactive tabletop. However, large tabletops are costly and need space, raising the question to what extent the physical size of shared horizontal surfaces really pays off. In order to analyse the suitability of smaller-than-tabletop devices (e.g. tablets) as a low-cost alternative, we studied the effect of the size of a shared horizontal interactive workspace on users’ attention, awareness, and efficiency during cross-device collaboration. In our study, 15 groups of two users executed a sensemaking task with two personal tablets (9.7’’), and a horizontal shared display of varying sizes (10.6’’, 27’’, and 55’’). Our findings show that different sizes lead to differences in participants’ interaction with the tabletop and in the groups’ communication styles. To our own surprise we found that larger tabletops do not necessarily improve collaboration or sensemaking results, because they can divert users’ attention away from their collaborators and towards the shared display.

Author Keywords
cross-device interaction; group work; tabletops; tablets; display size

ACM Classification Keywords
H.5.3 Information Interfaces and Presentation (e.g. HCI): Collaborative Computing

INTRODUCTION
Tabletops can greatly benefit collaborative activities such as visual analytics, search, and sensemaking [6, 9, 13], mainly because of enabling natural face-to-face communication around a shared workspace [7], providing a high level of awareness and fluid interaction [5, 19], and supporting natural territoriality during collaboration [21]. However, a large tabletop is also costly and difficult to maintain compared to other common off-the-shelf devices such as tablets. This is particularly true for environments like libraries or classrooms in which tabletops are often unavailable due to budget and space restrictions whereas tablets are quite common. Some researchers therefore aim at natural cross-device interaction and sensemaking with mobile devices instead of tabletops [10, 4, 17], while Tang et al. [22] recommend combining mobile devices with tabletops. Such combined cross-device systems (e.g. [11, 14, 18, 23]) are thought to be helpful for mixed-focus collaboration [3, 22] where users frequently transition from tightly-coupled collaboration to loosely-coupled parallel work, e.g., for reading or annotating a document or reviewing a long list of items on their own. Indeed, Wallace et al. [23] observed that in such scenarios the presence of a tabletop (39’’) leads to better sensemaking performance than with tablets alone. To better understand why, we closely studied the effect that three tabletops with different sizes had on users’ attention, awareness, and efficiency during cross-device mixed-focus collaboration. Investigating these effects enabled us to understand size-related advantages and disadvantages and to compare large tabletops with lower-cost,
smaller-than-tabletop devices that could potentially serve as alternatives. We conducted an experimental study of a collaborative sensemaking task that was executed by 15 groups of two users with two personal tablets (9.7”) and a shared workspace. We varied the size of the shared workspace over three conditions corresponding to off-the-shelf devices (10.6” tablet, 27” portable tabletop, and 55” large tabletop). To observe authentic collaboration during the study, we built a collaborative multi-device application for solving a complex analytical task based on the “Stegosaurus” data set from the VAST 2006 Challenge [16].

Our goal for the study was to better understand the implications of the size of the shared space and to contribute a nuanced description of its role in collaborative sensemaking activities. We address this overall research question by investigating the following subquestions both qualitatively and quantitatively:

- **RQ1:** How does the size of the tabletop affect users’ interaction with the devices?
- **RQ2:** How does the size of the tabletop affect the communication between the participants?
- **RQ3:** Does the size of the tabletop affect group awareness, the quality of the results of the group work, and participants’ satisfaction with the support by the system?

**RELATED WORK**

There are multiple studies in HCI and CSCW of how different physical affordances of interactive technology affect co-located collaboration, e.g., studies examining different tabletop sizes [20], different orientations (vertical vs. horizontal) [19], or different input modalities (touch vs. mouse) [5]. These studies pursued the goal of understanding how the physical affordances of a technology affect the kinds of cognitive and social interactions that will result, e.g., accessibility, visibility, ‘sharability’, etc. [19]. Our study follows this tradition of research and shares the goal of studying the effect of different affordances on co-located collaboration.

**Understanding Affordances & Tabletop Collaboration**

Rogers and Lindley [19] compared collaborative problem-solving around horizontal tabletops with collaboration at large vertical screens. They found that tabletops made it easier for all members to contribute to the problem-solving, encouraged group members to swap roles more often, and to explore more alternative solutions. Ryall et al. [20] compared collaborative word-puzzle solving with different tabletop sizes (80 cm vs. 107 cm diagonal) and group sizes (2-4 members). Their study revealed that users preferred a large tabletop for this particular task, but that size had no effect on completion time and the distribution of work. Interestingly, Ryall et al. mention that examining size-related issues for multi-device and mixed-surface interaction should be part of future work.

Morris et al. [12] provide a survey of collaborative search on tabletops and identify benefits of tabletops such as face-to-face collaboration that encourages more equal working styles than shoulder-to-shoulder situations. However, they mention that orientation of content on tabletops is an issue that could be solved with additional devices like tablets that enable e.g. text input or reading of documents.

Tang et al. [22] analyzed collaboration in tabletop settings with groups of two. In two studies they identified different coupling styles during mixed-focus collaboration [3] that range from closely coupled collaborative situations to loosely coupled parallel work activities. To support the latter, Tang et al. recommend providing mobile high resolution personal territories (e.g. tablet PCs or PDAs) as an addition to the horizontal shared surface of the tabletop.

This related work on physical affordances and size-related issues for multi-device/cross-device interaction serves as a motivation for our study and for better understanding the role of tabletops of different sizes in mixed-focus collaboration.

**Tablets and Tabletops**

With the growing popularity of cross-device and multi-device interaction, researchers have proposed working with multiple mobile devices instead of a tabletop. Lucero et al. [10] have introduced several cross-device mobile applications for seamless collaboration across smartphones, e.g., for photo sharing. Hamilton and Wigdor [4] propose Conductor, a system facilitating multiple tablets for single-user analytical sensemaking. Like we did in our study, they used the VAST 2006 data set to study the use of their system. They found that users used between 5-10 tablets for their task and that the parallel and combined use of the tablets was found to be highly useful.

Other systems have explored a close integration of mobile interaction with tabletops, e.g., for visual exploration [11], for sensemaking [23], for discussing and annotating work results in a science lab [14], or for collaborative Web search in schools [18]. McGrath et al.’s Branch-Explore-Merge system uses a private view on a tablet and a public view on a tabletop for collaborative visual exploration [11]. Private and public views are coupled using a branch-explore-merge protocol similar to revision control systems for source code, so that the results of individual explorations on the tablets can be merged back into the group results at any time. Wallace et al.’s prototype system was used to study collaborative sensemaking with only tablets, only a tabletop, or a combination of both [23]. They found that the presence of a tabletop led to improved sensemaking performance and supported a group’s ability to prioritize information, to make comparisons between task data, and to form and critique the group’s working hypothesis. They also found a positive correlation with equity of member participation using the shared digital table and a negative correlation of equity of member participation using personal tablets. Albeit similarities in system functionality like document transfer from tablet to tabletop and vice versa, the key difference between their work and our investigation is the variation of tabletop size as an independent variable as compared to tabletop being present or not.

This related work, in particular Wallace et al. [23], inspired varying table size as an independent variable during a study. We hereby contribute a better understanding of the nuances of the role of the tabletop during collaborative sensemaking.
Studies Using VAST 2006 Challenge Data
HCI, CSCW, and InfoVis researchers have long been interested in understanding the benefit of tabletops or tablets for sensemaking. For this reason, many used the data and tasks from the VAST 2006 Challenge to observe authentic sensemaking activities during a controlled and sufficiently demanding task. The VAST 2006 Challenge was developed by the National Visual Analytics Center at Pacific National Labs and can be considered a standardized visual analytics task [16]. Andrews et al. [1] have used it to analyze how a large high-resolution display affects the use of space and working styles. Hamilton et al.’s tablet-based sensemaking system Conductor was also studied using VAST 2006 data [4]. Isenberg et al.’s Cambiera used the same data to study collaboration activities in tabletop settings for dyads and identified different coupling styles describing the groups’ activities [6]. Jakobsen and Hornbæk used VAST data on high-resolution wall displays to study co-located collaboration [8]. Their system resembled Cambiera and they found that a wall display also supports various coupling styles and can be used to observe them and their relation to the proximity of participants.

In the light of this extensive use of the data and tasks of the VAST 2006 Challenge for studying (collaborative) sensemaking, we also chose to use them for our study.

STUDY
In order to investigate the effect of tabletop size on collaborative cross-device sensemaking, we designed and implemented a multi-device system that enabled groups to search, read, and make sense of documents from a database. The system served as experimental platform for observing (close to) authentic collaborative behavior. Obviously, the use of a specific system with a particular interaction design influences user behavior and thus restricts the extent to which we can generalize from our findings to all collaborative tasks and systems. We discuss such limitations in greater detail later.

Pre-Study
Our research questions, our study design, and the system design were informed by a pre-study. It was conducted with an earlier version of our system, where 15 pairs were asked to work on a general search-related group task. Already the very initial analysis of the collected video data indicated that the efficiency gain from a larger display was far smaller than expected – or even negative. In order to look into this phenomenon more systematically, we refined our system for the main study and focused the UI and study design on our identified sub-topics: the interaction with the devices (RQ1), the communication activities (RQ2), and the quality of group work results and users’ satisfaction with system support (RQ3).

Task
To provide a documented, authentic, and sufficiently demanding collaborative sensemaking task, we used the “Stegosaurus” data from the VAST 2006 Challenge [16] that was already used by [1, 4, 6, 8]. VAST 2006 describes a fictional story with 238 news articles, images, and data sheets. Groups have to search and find relevant information from this data, filter non-important articles, and connect important facts that lead to new insights and suppositions in order to come up with a solution to a hidden plot. The task involves solitary activities like searching for and reading relevant information as well as group activities like sharing and sorting relevant documents in order to solve the case. The task is difficult enough to require participants to work together for an extended period of time (we limited it to 90 minutes).

We did not expect groups to completely solve the task within the time frame. However, participants were handed questions about the plot at the beginning of each session to focus their behavior and stimulate an authentic collaboration.

System Description
Our system supported groups during phases of solitary and collaborative work. On their tablets, participants could execute full text searches in the VAST 2006 data by typing one or more keywords into a search box. Search results were presented in a ranked list (right tablet in Figure 2). Each item in this list represented a document and contained a document ID for clear verbal reference, a preview of sentences that contained the keyword(s), and icons indicating whether a document was already read, favored, or edited before. To access the document users tapped on the respective search list item to open a document view (left tablet in Figure 2). This view enabled them to read and highlight text in a document, add comments to it, share it on the tabletop, or send it directly to the tablet of their group partner on which the document then was also displayed in a document view. Each participant was assigned with an individual color (green or yellow). This color-coding was used for the background-color of their tablet, as the highlight-color of selected parts of documents, and as an indicator for the origin of each shared document.

Figure 2. Overview of the system with two tablets and one tabletop.

Shared documents appeared on the tabletop in boxes containing the document ID, an excerpt of its content, and buttons for document-related actions (tabletop in center of Figure 2). The possible actions enabled participants to reread documents on their personal tablet or to remove documents from the tabletop. Each box was additionally marked with a bookmark icon with a participant’s color to indicate who shared the document. Boxes could be moved and rotated freely on the tabletop to allow for spatial arrangements or to form tableaux to
embody the groups working hypotheses [23]. Multiple boxes could be grouped into visual clusters using touch input to encircle them with a lasso. This created a colored convex hull around the selected boxes as a visual grouping. This hull remained around the boxes and its shape was updated whenever the boxes were rearranged by participants. For a rapid exchange of information users could send documents to the tablet of the other group member. Adding documents to the tabletop enabled users to externalize ideas, store documents for later use and to create a common ground for group communication. Clustering of snippets helped to find connections between relevant events, persons, and places. The color-coding provided a visual connection between the participants, their tablets, and their contributions to the tabletop.

**Apparatus**

Figure 1a shows the study setup, which consisted of a conventional office table (1.4 × 0.8 m) on which the tabletop that served as the shared space was placed. For the study, we chose three multi-touch devices of different sizes as tabletops (Figure 1b, c, and d): 10.6′′ (Microsoft Surface 2 Pro), 27′′ (Lenovo Horizon), and 55′′ (Microsoft Perceptive Pixel). Each of them represented a typical size for a personal mobile device, a personal workspace, and a multi-user workspace. All three devices had the same display resolution of 1920 × 1080 pixels to guarantee for a higher internal validity of the study.

The two participants sat on office chairs and were facing each other. Each one was provided with an Apple iPad 2 (9.7′′) as their tablet. Next to each participant was a small mobile desk for filling out the questionnaire and as a space for notes or to place the tablet.

We used two cameras to record the study. One camera was installed with a bird’s-eye view above the table (Figure 1 right) and a second camera recorded the scene from the side to be able to see gestures, postures, and interactions of both participants (Figure 1 left).

**Participants**

30 participants (15 female, 15 male) were recruited for the experiment. The mean age was 22.4 years (SD = 2.9, aged 15-28). 29 participants were right-handed and one left-handed. None of the participants had color vision deficiency, consequently they had no problems with the employed color-coding. 25 participants were undergraduate students and 4 participants PhD students from non-technical subjects such as psychology, biology, or politics. One participant did not answer this question. To recruit the participants we used postings and flyers looking for “detectives”. By this we hoped to find participants who were interested in puzzles, crime stories, or mysteries and thus were motivated by the VAST 2006 system. We employed several data collection methods in order to investigate our research question: We gathered video data in order to analyze participants’ group work behavior, data logs in order to analyze and quantify participants’ interactions with the included digital devices, and a questionnaire and interviews in order to investigate participants’ results of group work and to what extent participants felt supported by the system.

**Video Data**

To address the subtopics of our research questions, we developed a coding scheme to analyze our video data. Doing so, we placed the focus on participants’ interaction with devices, their communication behavior, as well as the object they focused on during the task. After initial viewings of our videos, we started to describe and distinguish between different codes that apply to our research questions as well as offering a sound description of the video data. This coding scheme was refined until we felt that it was saturated. After this, an inter-coder reliability test was conducted for three 10 minute samples of video data to ensure objective coding by our two independent coders. Cohen’s Kappa was calculated individually for the categories Focus (κ = 0.96), Communication (κ = 0.85), and Interaction (κ = 1.0), revealing a high inter-coder reliability in each category for the final coding scheme that is presented below.

**Procedure**

We chose a between-subjects design with the tabletop size as independent variable with three conditions (10.6′′, 27′′, 55′′). In total, we investigated five pairs in each condition resulting in 15 pairs and 30 participants respectively. Like Isenberg et al., we chose a group size of two, because in their study two group members working on VAST 2006 were already sufficient to identify differences in coupling styles and collaboration strategies [6].

At the beginning of each study participants were seated at the two mobile desks. They were asked to fill out a questionnaire about demographics and tech-savviness. Then they were seated facing each other on the long sides of the table. Participants received an introduction into all features of the system and were given time to explore its functionality with a sample data set until they felt comfortable enough using it. The experimenter then introduced them to the VAST 2006 task. During the 90 min. session, they could look up task instructions at any time on their tablet or on a printed sheet of paper. The experimenter answered general questions about the procedure or prototype but did not answer questions about participants’ performance or how to solve the given task. Participants were free to move the mobile desks and office chairs as desired. The session stopped after 90 min or when participants felt that they had solved the case and found the plot. Then, both group members were asked to individually fill out a second questionnaire about their findings about the plot. Finally, we conducted a short semi-structured interview in order to ask participants about the strategy that they employed during the task and how much they felt that the system supported their activities. Each session lasted about 2 hours in total, and afterwards participants were compensated for their time.

**DATA ANALYSIS**

We employed several data collection methods in order to investigate our research question: We gathered video data in order to analyze participants’ group work behavior, data logs in order to analyze and quantify participants’ interactions with the included digital devices, and a questionnaire and interviews in order to investigate participants’ results of group work and to what extent participants felt supported by the system.
0.0 Focus: In order to investigate where participants placed their visual focus during the task, we distinguish between the codes 0.1 Tablet, 0.2 tabletop, 0.3 Person and 0.4 Other

1.0 Communication: To analyze the extent to which participants were talking, we code 1.1 Talk and 1.2 Silence. Utterances and mufflings were coded as 1.2 Silence.

2.0 Interaction: To investigate participants’ interaction with the tabletop, we code 2.1 add document for sending documents from the tablet to the tabletop, 2.2 delete document for deleting documents from the tabletop, 2.3 move document for moving documents to another place on the tabletop as well as 2.4 cluster documents for participants’ activities to collect documents in a cluster in order to sort them.

We used the coding scheme for both qualitative and quantitative investigation. In our qualitative analysis, the creation of the coding scheme helped us to elicit the main aspects of the group work. Also, it allowed us to focus on specific subparts of the video for a more detailed observation. We could, for example, only view sequences when participants where focusing on the table in order to further investigate this event by describing its context and related events in more detail. In addition, we also calculated the frequencies and durations for the coded events in order to investigate how the qualitative findings are reflected in the quantitative data.

Data Logs
In addition to the video data, we also collected logs of users’ interaction with the system. We logged the number of document manipulations (e.g. move) on the tabletop and how often documents were transferred from the tabletop to a tablet or vice versa. We also logged the number of times a document was transferred from one tablet to another. We did this in order to analyze whether the tabletop size had an effect on how and how often users interacted with documents.

Questionnaire and Interview
We used a questionnaire with questions about the solution (i.e. plot) of the VAST 2006 challenge to understand (1) the quality of the group results based on individual answers to the questions and (2) the extent to which group members had the same knowledge of their results and solutions to the task. Also, we concluded each session with a semi-structured interview focusing on group strategies when working on the task and to what extent participants felt supported by the system.

FINDINGS & DISCUSSION
In this section we report on our findings and discuss them in relation to our research questions. In particular, we focus on the differences across the three different sizes of the tabletop (SMALL: 10.6”, MEDIUM: 27”, LARGE: 55”) with respect to users’ interaction with the devices (RQ1), the group’s communication activities (RQ2), and the differences regarding user perceptions and the quality of group work results (RQ3). For each research question, we elaborate on specific subtopics based on our qualitative observations and interview data. Quotes from participants in the following section were slightly modified to prevent publishing the solutions of the VAST 2006 Challenge tasks.

In addition, we report our quantitative results in relation to our observations. In our analysis, we employed a mixed-effects models approach to analyze individual scores taking participants’ group membership as an additional source of variance into account. We used the nmlme package by Pinheiro and Bates [15] for R.

RQ1: Findings on Device Interaction
To answer RQ1, we distinguish between three different aspects of the participants’ interactions with devices that could be influenced by the size of the tabletop: (1) participants’ visual focus and attention, (2) participants’ frequency of searching and transferring documents, and (3) participants’ manipulation of documents on the tabletop.

Focus and Attention
In our qualitative analysis of video data, we observed that participants used the system mostly as we expected and intended in all three conditions. They used the tablets for searching or reading activities and the tabletop for collecting and grouping documents. But we also observed the following differences between the conditions: Participants in the SMALL condition used the tabletop in a more goal-oriented fashion and only focused on it when actively using it. In comparison, participants in the LARGE condition stayed visually focused on the tabletop even during phases of reflection or for visually scanning over shared documents. Interestingly, this was notable even during phases that did not require visual focus on the tabletop (e.g. when reflecting about next steps) but during which the tabletop apparently attracted most participants’ visual attention. For example, participants sometimes looked at and “played” with documents on the tabletop (i.e. playfully moving them back and forth with their fingers without the intention of actually reordering them) while thinking about how to proceed with the task and even while verbally coordinating next steps with their partner (“What could we search for that we don’t already have on the table?”) or reflecting on the progress so far (“What do have we already? How is all this related?”).

In contrast, participants in the SMALL condition were visually focusing on and directly referring to their group partner when they were reflecting or coordinating their activities (e.g. one participant who was finished reading a document on his tablet looked up and directly addressed his group partner: “Did you find anything yet?”). These observations are also reflected in our quantitative results. Table 1 shows the average percentage of the session time that participants focused on the tablet, tabletop, group member, and other objects across the three conditions.

Our analysis revealed no overall statistically significant effect of “focus on other person” ($F(2,12) = 2.74, p = .10$). But as Table 1 shows, participants using the small tabletop focused three times longer on the other person (14% of the time) than participants using the large tabletop (4% of the time). With a statistically significant difference between LARGE and SMALL ($b = 9.77, t(12) = 2.22, p < .05$). There were, however, no statistically significant differences for this
between the SMALL and MEDIUM conditions and no statistically significant differences between the MEDIUM and LARGE conditions.

The analysis also shows statistically significant differences for “focus on tabletop” (F(2,12) = 6.50, p < .05) between LARGE and SMALL (b = -3.17, t(12) = -2.63, p < .05) and LARGE and MEDIUM (b = -3.88, t(12) = -3.21, p < .05). Participants using the LARGE tabletop focused almost twice as long on the tabletop (10% of the time) than in the MEDIUM condition (6% of the time). There were, however, no statistically significant differences for this between the SMALL and MEDIUM conditions. Focusing on other things (e.g. on the information sheet or just looking around) happened to the same extent in all conditions.

Finding 1.1. Our results show that participants’ visual attention in cross-device settings can be drawn towards a large tabletop, even in situations where there is no pragmatic reason to focus on the tabletop. This can result in non-“face-to-face” verbal communication without eye contact between collaborators. On the contrary, a small tabletop received less visual attention and was used in a more goal-oriented fashion. Users did not “play” with objects on the small tabletop and focused significantly longer on their collaborator.

The choice of the size of a tabletop should therefore not been taken too lightly. Larger does not necessarily mean better. Depending on the context of use, it can be desirable to increase eye contact within a group, (e.g., to build trust in collaborators or creating a sense of intimacy) or to increase the focus on the shared tabletop (e.g. to support collaborative activities).

Search and Transfer

Our qualitative data revealed differences in the strategies that participants employed to access documents across the conditions. We observed that participants worked more filing-oriented in the LARGE condition and more search-oriented in the SMALL condition.

By filing-oriented we mean that the tabletop served as a space to collect, file, and pile found documents in space and to use visual and spatial cues to quickly access and then transfer documents to the tablets. We observed that the large tabletop display was used as external memory and “space to think” as described in [1].

By search-oriented we mean that even after a found document was filed or piled on the tabletop, participants frequently accessed these previously found documents by repeated searches with the tablet using memorized keywords rather than looking up documents on the tabletop and sending them to the tablets.

We found evidence for these two different strategies in our video data (e.g. one user commented “I will just search for ‘basement’ - this will get me the document about the groundskeeper.” Another user asked “Where did we put the documents with the wine export?” while trying to find a document on the tablet). However, we could not directly confirm these findings with a quantitative analysis. For this analysis we attempted to use the number of document transfers across devices as indicators for the two strategies. We distinguished between (1) the transfer of documents from a tablet to the tabletop, (2) the transfer of documents from the tabletop to a tablet, and (3) the transfer of documents between tablets. We expected that a filing-oriented strategy should lead to more frequent tabletop-to-tablet transfers than a search-oriented strategy.

Table 2 shows the mean number of document transfers per participant.

Table 2. Document transfers across devices per participant.

<table>
<thead>
<tr>
<th></th>
<th>Tablet to Tablet</th>
<th>Tablet to Tabletop</th>
<th>Tabletop to Tablet</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>3.3 (SD=4.76)</td>
<td>6.2 (SD=2.94)</td>
<td>12.6 (SD=8.06)</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>5.1 (SD=5.09)</td>
<td>8.5 (SD=3.74)</td>
<td>14.3 (SD=10.41)</td>
</tr>
<tr>
<td>LARGE</td>
<td>1.3 (SD=1.25)</td>
<td>6.5 (SD=3.21)</td>
<td>23.7 (SD=12.6)</td>
</tr>
</tbody>
</table>

Finding 1.2. There was qualitative evidence in our data that a larger tabletop affords more spatial and visual organization schemes for digital content, which is also considered one of the key qualities of large high-resolution displays [1] or tabletops [9] for sensemaking.

Nonetheless, we observed how users skillfully used the search function and memorized keywords to successfully compensate for the lack of space on a medium-size or small tabletop. This is noteworthy, since (as we discuss later) there was no significant difference in the objective quality of group results for the different tabletop sizes.
organization of documents. In contrast, participants in the LARGE condition did not always interact purposefully with documents but tended to “play” with documents by moving them on the tabletop. Participants also spent a lot of time browsing and arranging the content of the tabletop while seemingly getting sidetracked from the task at hand. This could imply that the larger tabletop affords a more spatial, but also more playful and serendipitous interaction with the documents, while the smaller tabletop affords a limited but more goal-oriented interaction.

In Table 3, we summarize our quantitative analysis of participants document manipulation on the tabletop. It shows the mean number of different manipulations (add, delete, move, cluster) per session.

The analysis shows how tabletop size did indeed affect the number of document movements on the tabletop. It revealed a significant effect of the number of move-manipulations ($F(2,12) = 3.98, p < .05$) between LARGE and SMALL ($b = -12.0, t(12) = -2.32, p < .05$) and LARGE and MEDIUM ($b = -12.0, t(12) = -2.32, p < .05$).

Participants in the LARGE condition moved documents almost twice as often as participants working with MEDIUM or SMALL while there were no significant differences in add, delete, or cluster. This hints at a more exploratory working style with frequent rearrangements of documents and a more playful style of interaction.

**Finding 1.3.** Based on our qualitative and quantitative analysis, we found that a larger tabletop can invite a more explorative or playful interaction with objects on the tabletop that is also reflected in a significantly higher number of movements. Other kinds of manipulations (e.g. clustering) showed no statistically significant differences.

This more frequent movement of objects is not necessarily connected to purposeful rearrangement of objects in space for sensemaking. Instead, larger tabletops seem also to invite more playful and serendipitous interactions that happen between and during the actual goal-oriented manipulations or conversations.

**Amount of Communication**

Our qualitative observations indicate that the amount of communication differs across the three conditions. The larger devices seemed to attract and hold attention and reduce verbal communication.

For example, we observed that participants in the SMALL condition often explicitly dedicated time to engage in communication with their group partner. Doing so, they would lean back in their chairs and shift their attention to the other participant and away from the devices. This “leaning back” was intended for signaling readiness to communicate to the other group member. Typically this was followed by direct verbal communication about the current results and the next steps. It occurred during group work, when participants stepped out of their current activity in order to reflect on the work done so far and think about next steps. Thus it can be regarded as an active measure to create group awareness.

In the LARGE condition, participants only rarely took the time to directly address their group partner without focusing on the tabletop. They reflected on their work not by directly addressing their group partner but through activities mediated by the tabletop such as browsing and rearranging the content on the tabletop. This was not necessarily accompanied by verbal communication.

**Finding 2.1.** Based on our qualitative analysis, we found that participants talked more and more directly to each other when using a small tabletop than when using a large tabletop. We observed that this face-to-face talk seemed to be more important when using a smaller tabletop and that participants introduced phases exclusively dedicated to verbal communication for coordinating actions and creating awareness. This suggests that verbal communication was used to compensate for a smaller shared space that provided less overview and only allowed for less effective deictic references. However, a quantitative analysis did not reveal a statistically significant difference in the total amount of communication.

**Table 3. Mean number of different manipulations of documents on the tabletop per session. Statistically significant differences ($p<.05$) are shown in bold.**

<table>
<thead>
<tr>
<th></th>
<th>Add</th>
<th>Delete</th>
<th>Move</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>5.1</td>
<td>0.6</td>
<td>9.7</td>
<td>4.1</td>
</tr>
<tr>
<td>(SD=2.43)</td>
<td>(SD=0.80)</td>
<td>(SD=6.12)</td>
<td>(SD=2.84)</td>
<td></td>
</tr>
<tr>
<td>MEDIUM</td>
<td>5.8</td>
<td>1.5</td>
<td>9.7</td>
<td>2.5</td>
</tr>
<tr>
<td>(SD=3.74)</td>
<td>(SD=1.20)</td>
<td>(SD=6.72)</td>
<td>(SD=3.26)</td>
<td></td>
</tr>
<tr>
<td>LARGE</td>
<td>5.8</td>
<td>0.9</td>
<td>21.7</td>
<td>2.8</td>
</tr>
<tr>
<td>(SD=2.75)</td>
<td>(SD=0.94)</td>
<td>(SD=14.51)</td>
<td>(SD=1.72)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4. Percentage of time that participants were talking to each other.**

<table>
<thead>
<tr>
<th></th>
<th>Talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>16.51</td>
</tr>
<tr>
<td>(SD=9.03)</td>
<td></td>
</tr>
<tr>
<td>MEDIUM</td>
<td>13.68</td>
</tr>
<tr>
<td>(SD=4.95)</td>
<td></td>
</tr>
<tr>
<td>LARGE</td>
<td>8.39</td>
</tr>
<tr>
<td>(SD=4.68)</td>
<td></td>
</tr>
</tbody>
</table>

**RQ2: Findings on Communication Activities**

Our analysis differentiates between the amount of communication and characteristics of communication. Both are discussed in the following.
Characteristics of Communication

In order to investigate what mediates communication across the three conditions, we analyzed where participants were looking while they were talking to each other.

Based on our video observations, we also found differences in communication with respect to how participants used deictic references by pointing to documents on the tabletop for elaborating their arguments. For example, participants working with the large tabletop were often already engaged in sense-making discussions while they were interacting with the documents on the tabletop. This was also reflected in the prevalence of deictic references that accompanied their communication. For instance, “What do we have here? We have the groundskeeper cleaning the basement (participant pointing at the tabletop), we have the event at the town hall (pointing), and the sick grandparents (pointing). How could these three be related?” In the MEDIUM and LARGE condition the two group work activities structuring documents on the tabletop and talking about and making sense of the gathered information blended into each other and participants frequently switched between them. When group partners talked during such activities, this happened often without eye-contact between the collaborators and was mediated by the tabletop.

In contrast, the two activities structuring documents on the tabletop and talking about and making sense of the gathered information rarely overlapped in the SMALL condition. After participants clustered the documents on the tabletop according to document topic, they would disengage from the tabletop and address their group partner in order to reflect on the information. For example, one participant was silently clustering documents on the tabletop while the other participant was reading a document on his tablet. Once the participant was finished with his sorting activity, he withdrew from the tabletop, rubbing his face, and addressing his partner by saying: “Let’s think about what we have so far,” upon which his partner looked up and they started a discussion about the information that they have read and found so far without using deictic references.

Our quantitative results confirm these differences for the visual focus. Table 5 shows the average percentage of the session time that participants focused on the tablet, tabletop, group member, and other objects while talking. Results indicate that participants mostly focused on their tablet while talking. This is not surprising and lies in the nature of the task, in which the tablets were the source of the information that participants wanted to convey to their partner. For instance, they looked at it to read important information aloud or to verify that their contribution was correct.

Our analysis revealed a significant effect on the time participants focused on the tabletop while talking ($F(2,12) = 8.03$, $p < .05$) for LARGE and SMALL ($b = -13.08$, $t(12) = -3.39$, $p < .05$) and LARGE and MEDIUM ($b = -12.27$, $t(12) = 3.18$, $p < .05$). In the LARGE condition, participants focused almost twice as long on the tabletop (24.6%) than in the MEDIUM (12.33%) or SMALL condition (11.53%). In addition, Table 5 also shows that participants in the SMALL condition focused more on the other person while talking, whereas participants in the LARGE condition focused more on the tabletop than on the other person.

Table 5. Individuals’ focus while talking. Statistically significant differences ($p < .05$) are shown in bold.

<table>
<thead>
<tr>
<th></th>
<th>Talk &amp; Focus</th>
<th>Talk &amp; Focus</th>
<th>Talk &amp; Focus</th>
<th>Talk &amp; Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tablet</td>
<td>Tabletop</td>
<td>Person</td>
<td>Other</td>
</tr>
<tr>
<td>SMALL</td>
<td>49.27 (SD=10.13)</td>
<td>11.53 (SD=6.65)</td>
<td>29.76 (SD=15.41)</td>
<td>9.44 (SD=4.66)</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>53.73 (SD=11.53)</td>
<td>12.33 (SD=8.08)</td>
<td>24.96 (SD=11.92)</td>
<td>8.98 (SD=5.07)</td>
</tr>
<tr>
<td>LARGE</td>
<td>49.36 (SD=14.84)</td>
<td>24.61 (SD=8.88)</td>
<td>18.11 (SD=7.60)</td>
<td>7.92 (SD=5.00)</td>
</tr>
</tbody>
</table>

**Finding 2.2.** Our data shows that the large tabletop attracted the visual attention of group members during verbal communication. It also enabled a close integration of actively structuring and manipulating items on the tabletop with talking about and making sense of the items including the use of deictic references.

This close integration did not happen for the small tabletop where active structuring and talking rarely overlapped. Instead these activities happened sequentially with frequent phases that were exclusively dedicated to verbal face-to-face communication (including eye contact) without interacting with the tabletop.

**RQ3: Findings on User Perceptions and Group Results**

We analyzed data from our interviews and questionnaires to understand (1) participants’ collaboration strategies during group work, (2) participants’ satisfaction with the system support, and (3) the objective quality of the results of group work.

**Group Strategy and Awareness**

Concerning collaboration strategies, participants in the SMALL condition reported that they had not explicitly distributed tasks among themselves. Rather they worked mostly solitarily while keeping their group partner informed about their current progress. They also stated that most of the time they were aware of their group partners’ activities and that “there was no need to communicate the strategy – only findings needed to be communicated.” They stated that their main focus was on reading and that clustering and sorting on the small tabletop was considered as taking “too much time.”

Participants in the MEDIUM condition reported that they occasionally distributed tasks among each other, but that a more explicit assignment of subtasks would have improved group work. Despite the division of labor, they were mostly aware of their partner’s activities, but they sometimes needed to ask their partner for a progress update (“What are you currently working on?”) to actively create group awareness.

In the LARGE condition, participants stated that they frequently distributed tasks among themselves. They also reported that they were only partially aware about their group partner’s activities, but some awareness was mediated by the tabletop. Clustering and sorting activities were perceived to help them in sensemaking and reflection. One participant said: “If I didn’t know how to proceed, I just looked at the table.”
User Satisfaction with System Support

Participants in the SMALL condition valued the support for direct document transfer between tablets and the highlighting function to emphasize important parts of a document. They were also satisfied with the ability to cluster documents. Other advantages of the tabletop were not mentioned. They reported that they did not divide the tabletop into personal and group territories [21], but, in contrast, explicitly valued the clear distinction between personal activities on the tablet and shared activity on the tabletop (as suggested by Tang et al. [22]).

Participants in the MEDIUM condition mainly valued the tabletop as an information and memory pool that allowed them to quickly access relevant documents. Furthermore, they reported using the space in order to sort documents thematically. Most of them stated that they were satisfied with the size of the table, however some mentioned that a larger tabletop might have allowed for additional personal spaces on the interactive table.

Finally, participants in the LARGE condition explicitly stated their appreciation of the tabletop as an efficient but also playful way of accessing documents, while providing a good overview. They additionally stated that they divided the tabletop into different areas that they used to distinguish between kinds of documents. For example, one group reported that they “stuck all the unimportant stuff in the corner and focused on the main documents of the plot in the center of the table” which resonates with the observations of Andrews et al. about sensemaking on large screens [1]. Participants used the tabletop to structure documents and to compare them with documents displayed on the tablets. Furthermore, they valued being able to work on the tabletop simultaneously and reported the interaction with the system to be fun. Regarding the size of the display, participants were satisfied and positively reported that it even allowed them to have a personal territory on the shared table as well. Some of the participants expressed a wish to also have a reading function on the tabletop in addition to that on the tablet.

Results of Group Work

To assess the results of the participants’ group work we asked them to fill out a questionnaire about facts on the plot they were investigating. This questionnaire consisted of nine multiple-choice questions. In total, participants were able to get 44 points, depending on their answers. Table 6 summarizes participants’ general performance across the three conditions as well as the level to which they agreed on an answer irrespective of whether it was right or wrong. For the Performance Score significance was tested based on individual scores using the mixed-effects model approach. For the Conformity Score we applied a one-way ANOVA based on group scores that show how consistent the group partners performed.

The results show no statistically significant differences for the Performance Score or for the Conformity Score. This lets us conclude that in all three conditions the participants were provided with sufficient support and flexibility to establish successful collaboration styles and strategies. Neither the system nor the task unfairly privileged one condition compared to another.

The qualitative analysis of how participants made use of the different tabletop sizes confirms the benefit that is usually attributed to large displays or tabletops for increased awareness [13, 19] or for spatial sensemaking [1, 6]. The appreciation of the large tabletop by participants and the desire for a larger tabletop by the participants in the MEDIUM condition fit into this picture and resonate with earlier findings about the desire for more space during around-the-table collaboration [20]. The positive attitude towards a clear separation between private territories on mobile devices vs. group territories on tabletops is in accordance with earlier work on territories and coupling styles around tabletops [21, 22].

However, we found it surprising how the subjective preference for larger tabletops was not reflected in the objective quality of the results of the group work. While there were clear differences in group awareness and task distribution, there were no statistically significant differences in Performance and Conformity scores. This indicates that the participants were perfectly able to perform sensemaking and maintaining awareness using face-to-face communication, rather than it being primarily mediated by the tabletop.

Strategies included the introduction of different styles of communication with eye-to-eye contact, using search-oriented strategies instead of filing-oriented strategies, and relying more on direct exchange of information between tablets than sharing documents via the tabletop.

Table 6. Mean Performance Scores (mixed-effects model, based on individual scores) and Levels of Conformity per group (one-way ANOVA, based on group scores).

<table>
<thead>
<tr>
<th></th>
<th>Performance Score</th>
<th>Conformity Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>83% (36.7 / 44 - SD 3.86)</td>
<td>81.8% (36.0 / 44 - SD 3.98)</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>76% (35.4 / 44 - SD 3.67)</td>
<td>73.6% (32.4 / 44 - SD 2.3)</td>
</tr>
<tr>
<td>LARGE</td>
<td>76% (33.3 / 44 - SD 3.16)</td>
<td>76.8% (33.8 / 44 - SD 3.27)</td>
</tr>
</tbody>
</table>

Finding 3. From our observations we can conclude that the users working with small tabletops during cross-device collaboration were perfectly capable of maintaining awareness and performing sensemaking primarily mediated by face-to-face communication. They were able to adapt and improvise and succeeded in achieving an objective quality of work results that was not significantly different from those working with medium-sized or large tabletops.

LIMITATIONS

As mentioned in the study description, the system design and the given task naturally influenced the collaboration behavior that we observed. There are specific aspects of our study that limit the generalizability of our findings to all cross-device systems for sensemaking: First, the VAST 2006 Challenge made it necessary to frequently search for documents and
to closely read them. This shifted the overall nature of the collaboration towards many phases of solitary parallel work and fewer phases of tightly-coupled collaboration. Users focused on their tablets at least for 71% of the time (Table 1). A different task (e.g. collaborative design or gaming) might afford more tightly-coupled collaboration and thus emphasize the role and the importance of the tabletop. This could also lead to a generally greater importance and benefit of larger tabletops than we observed. Second, our groups consisted of two users. We chose pairs like Isenberg et al. [6] who demonstrated that pairs are sufficient to identify differences in coupling styles and collaboration strategies. Choosing larger groups in followup studies might reveal a greater need for display space than in our study. In larger groups there will also be an increased importance of group awareness – whether it is actively created through verbal conversation or is mediated by the size and physical affordances of the tabletop. Third, our participants were students whose analysis skills differ from that of professional analysts. But as our study focused on observed differences between the three conditions and their UI and demands for technical proficiency were almost identical, our findings should be independent from education and technical skills of participants. Observed differences should not be substantially different for non-students. However, age-related differences in memory, eye-sight or motor skills could have an impact (better eye-sight and memory could benefit SMALL, motor impairments could hamper LARGE).

Nonetheless, we believe that our study employed a balanced task set and also realistic setting for collaboration. That none of the conditions can be considered unfair or unsolvable becomes evident through the absence of statistically significant differences in the groups’ performance scores.

IMPLICATIONS FOR FUTURE RESEARCH
Our findings point towards the need for a more nuanced understanding of the role of a shared tabletop and its size in co-located multi-device collaboration. We have observed that for the particular kind of collaborative sensemaking task that we studied, a larger tabletop did not lead to a better task performance or user experience. In particular, we found evidence that the large tabletop attracted significantly more attention than the smaller ones – for good and for ill.

However, a major part of our sensemaking task consisted of individual work with personal tablets, e.g., searching for documents and reading text. It would be interesting to contrast our results with a study of a task where the major part of the work requires tightly-coupled collaboration. It would also be interesting to see to what extent the found strengths of SMALL tabletop sizes (i.e. increased group awareness, focused work strategy) hold true for group sizes larger than two. Some effects that we considered to be positive in our study setting (e.g. focus on other person) might be detrimental to the quality of collaboration in larger groups.

Although we did not analyze territoriality in our study, findings suggest that the size of the tabletop influences how people manage personal and shared work spaces in cross-device settings. Research investigating the effect of shared display size on territoriality in group collaborations would further help to understand the role of the tabletop in relation to personal devices.

In addition, we are interested on how our findings relate to a study investigating the effect of a shared display size in a vertical setting. Changing the orientation of the shared display also implies changes to the physical user position and interaction with the system which might strongly influence the effects that we found. We would especially expect different results in how participants would refer to the shared display if it were vertically oriented.

Research on shared interactive surfaces has primarily focused on large tabletops or large interactive wall displays with up to 200” [2]. Our findings imply that under certain conditions small shared interactive surfaces could outperform large ones. While tablets are typically designed for personal tasks, our results encourage research on re-purposing tabletsized devices as shared interactive surfaces for mixed-focus and cross-device collaboration, or even to ad hoc expand the shared surface using additional devices (as proposed by [17]).

CONCLUSION
While interactive tabletops are increasingly combined with mobile personal devices such as tablets, the effect of size of such shared space is still unexplored. In this paper, we studied the effect of the size of a shared horizontal tabletop on users’ attention, awareness, and efficiency during cross-device mixed-focus collaboration. We conducted an experimental study of a collaborative sensemaking task using the “Stegosaurus” data set from the VAST 2006 Challenge [16]. We studied 15 groups of two users working with our collaborative multi-device application existing of two personal tablets (9.7”) and a tabletop with a variable size over three conditions (10.6”, 27”, and 55”). In detail, we focused on participants’ interaction with devices (RQ1), communication activities (RQ2) and user preferences and quality of work (RQ3). We found that larger tabletops do not necessarily improve collaboration or sensemaking results, because they can divert users’ attention away from their collaborators and towards the shared display. This has implications on participants’ interaction with the tabletop as well as on their communication style.

We conclude that the choice of size for a shared tabletop for cross-device collaboration is not obvious, but should depend on the task at hand, and that a larger tabletop is not necessarily the better choice.

ACKNOWLEDGEMENTS
We thank the German Research Foundation (DFG) for financial support within project C01 of SFB/Transregio 161.

REFERENCES


