SMAC: A Simplified Model of Attention and Capture in Multi-Device Desk-Centric Environments

ZHEN LI, University of Toronto, Canada
MICHELLE ANNETT, MishMashMakers, Canada
KEN HINCKLEY, Microsoft Research, United States
DANIEL WIGDOR, University of Toronto, Canada

Prior research has demonstrated that users are increasingly employing multiple devices during daily work. Currently, devices such as keyboards, cell phones, and tablets remain largely unaware of their role within a user's workflow. As a result, transitioning between devices is tedious, often to the degree that users are discouraged from taking full advantage of the devices they have within reach. This work explores the device ecologies used in desk-centric environments and complies the insights observed into SMAC, a simplified model of attention and capture that emphasizes the role of *user-device* proxemics, as mediated by hand placement, gaze, and relative body orientation, as well as *inter-device* proxemics. SMAC illustrates the potential of harnessing the rich, proxemic diversity that exists between users and their device ecologies, while also helping to organize and synthesize the growing body of literature on distributed user interfaces. An evaluation study using SMAC demonstrated that users could easily understand the tenants of user- and inter-device proxemics and found them to be valuable within their workflows.

CCS Concepts: • **Human-centered computing → Interaction paradigms**; *Interaction techniques.*

Additional Key Words and Phrases: Proxemics, user proxemics, inter-device proxemics, cross-device interaction, distributed user interfaces

ACM Reference Format:
Zhen Li, Michelle Annett, Ken Hinckley, and Daniel Wigdor. 2019. SMAC: A Simplified Model of Attention and Capture in Multi-Device Desk-Centric Environments. *Proc. ACM Hum.-Comput. Interact.* 3, EICS, Article 2 (June 2019), 47 pages. https://doi.org/10.1145/3300961

1 INTRODUCTION

The staggering adoption of mobile devices, such as tablets and smartphones, has enabled easy, on-demand access to digital content anytime, anywhere [28]. This prevalence also provides new opportunities for leveraging multiple physical devices to enable more fluid user workflows. A significant body of work has examined interaction methods (e.g., [11, 33, 42, 57, 66, 106]) and developed tools for building distributed user interfaces (e.g., [14, 24, 41, 62, 64, 68, 69, 105]). These projects demonstrated how multiple devices might blend seamlessly together and allow users to opportunistically choose preferred combinations of input and output technologies that are best suited to a given task. Absent from past work, however, has been a unified model of how consistent

Authors' addresses: Zhen Li, University of Toronto, Toronto, Ontario, Canada, zhen@dgp.toronto.edu; Michelle Annett, MishMashMakers, Newmarket, Ontario, Canada, michelle@mishmashmakers.com; Ken Hinckley, Microsoft Research, Redmond, Washington, United States, kenh@microsoft.com; Daniel Wigdor, University of Toronto, Toronto, Ontario, Canada, daniel@dgp.toronto.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2573-0142/2019/6-ART2 $15.00
https://doi.org/10.1145/3300961

user experiences may be developed and experienced across such devices, analogous to the inherent model of user attention which drives window focus and input device capture in traditional desktops.

A key macro-level challenge of cross-device interaction with modern commercial platforms is the lack of contextual information available to devices. As increasingly more users find value in form-factors such as tablets and smartphones, they have begun to integrate these devices into their everyday workflows. Prior work has found that many barriers exist when using such devices [78, 86]. Santosa and Wigdor [86], for example, found that while many users make extensive use of multiple devices, users generally worked on one device at a time due to the high cost of moving information between them. While Santosa and Wigdor's users expressed a clear desire for more cross-device fluidity, they relied on cloud services and email to pass information throughout their network of devices. In such scenarios, various applications are working in concert, *without an awareness* that they are being used as part of a larger workflow.

A primary mechanism for task-flow context available to an application on traditional computing platforms is the state of the window manager's inherent model of user attention. On desktop computers, for example, each application window knows whether it has the current attention or *focus* of the user. This window focus state composes one half of a simplified model of the user's attention, as described by Apple's Human Interface Guidelines: *"the foremost document or application window that is the focus of the user's attention is referred to as the main window"* [2]. The other half of this simplified model of attention is input capture, i.e., the application or user interface element that is the current target of pointer and keyboard input. For example, if a user depresses a menu item, it is presumed that the menu is now the focus of her attention, and it expands to cover-up other content in the view.

Each application window receives callbacks from the window manager and UI toolkit whenever the window focus or input capture changes. Because applications are typically composed of interface controls from a common toolkit, the responses to these callbacks are consistent across the entire system. For example, in both Windows 10 and Mac OS X, when a window loses the input capture of the keyboard, the cursor previously displayed in that window becomes hidden and the title bar of the application changes its visual appearance to reflect its lack of focus.

As each device within a multi-device ecosystem operates in relative isolation, with its own information about input capture and focus, there is no unified window manager or model of attention, unlike what is found on desktop computers. If a user stops looking at her phone and instead begins to type on her desktop system, for example, the foreground application on the phone receives no indication that it has lost the user's attention, nor does the desktop application receive any indication that it has gained the user's attention or that it had lost it while she was looking at the phone.

The goal of the present work is thus to propose a simplified model of attention and capture, or *SMAC*, for distributed user interfaces (DUIs). As a unified interaction model, SMAC is intended to sample the user's attention on different devices and route messages to the user based on the devices' proxemic relations, providing functionality similar to the system cursor and window state that drives the capture model of the Windows OS [70]. To develop SMAC, we reviewed past work on DUIs, including work focused on improving developer tools and user experiences, with a focus towards developing a model to support the desk-centric workflows found in past examinations of user behavior [86, 98]. From the many past projects, we focused on user experiences that leveraged the principle of proxemics [4, 29], which varied device behavior based on the user's physical posture with respect to the device. These past projects utilized an inherent model of user attention, and thus proved fertile ground for developing a unified model. As we examined the literature, we also identified that the proxemic relations *between devices themselves*, i.e., inter-device proxemics, was rarely-recognized as influential to understanding the attention of the user. Thus, the key concepts informing the dimensions of SMAC include user-device proxemic aspects that sample the user's attention such as the hand placement, gaze focus, and body orientation relative to computing devices, as well as inter-device proxemic aspects including the distance and the relative orientation between devices.

The contribution of this work is thus threefold. First, we contribute a unified interaction model, SMAC, that combines user- and inter-device proxemics, focusing on the interactions of a single user in a multi-device desk-centric scenario. SMAC should enable developers to create consistent, unified cross-device systems, and it also helps users better understand the behaviors of such a system. Second, we built OmniDesk, a prototype system with a collection of interaction techniques that demonstrates the utility of basing a DUI-based operating environment on the SMAC model. The utility of this prototype serves as validation of the completeness of the model for desk-centric tasks and should contribute to the development of interaction methods for distributed user interfaces. Third, we conducted an evaluation study and demonstrated that, after using OmniDesk to perform simple cross-device tasks, participants were able to understand and extrapolate the SMAC model to other multi-device activities. It is our intention that SMAC can serve as the DUI equivalent of the inherent model of user attention found in WIMP systems today.

2 RELATED WORK

The areas of cross-device interaction and proxemics have received much attention within the literature and are of most relevance to our exploration into user- and inter-device proxemics and the design of the SMAC model. Given the abundance of physical and digital information users have access to in their environments, prior work on information management also shaped the development of the SMAC model.

2.1 Cross-Device Interaction

The ubiquity and availability of mobile devices has generated a surge of interest into techniques that support cross-device interaction. Toolkits, frameworks, and connection techniques such as Conductor [33], SurfaceLink [27], XDBrowser [75], Weave [14], SyncTap [84], Pebbles [72], Tracko [45], CTAT [21], Augmented Surfaces [85], WatchConnect [41], Proximity Toolkit [64], PolyChrome [3], Panelrama [105], HuddleLamp [80], Orienteer [19] and Synchronous Gestures [37, 82], have explored different ways that information can be synchronized across devices. The reader is directed to Chong et al.'s [16] survey of interaction techniques for spontaneous device association for a detailed review of proposed approaches, many of which inspired the inter-device dimension of the SMAC model and interaction techniques implemented within the OmniDesk prototype system.

Many approaches for transferring data between devices involve the use of touch-screen gestures. Nacenta et al. [74] provides a useful survey of such techniques. "United Slates" [11] used a bimanual technique where the nondominant hand specifies the target and the dominant hand specifies the item to transfer. Marquardt et al. [66] employed unimanual gestures such as sliding and zooming that started on one device and terminated on another to share information. Lucero and colleagues [60] and Nielsen et al. [76] explored the use of pinching on mobile phones (i.e., one handed, two-handed, and two-stage) to share photos between phones. Rädle et al. [81] also utilized unimanual input but restricted it to a single device, harnessing spatially-agnostic or spatially-aware visualizations to transfer content between tablets. Duet [12] transferred sensor data between watch and phone based on the context of use in a given application scenario. ActivitySpace [42] explored activity-centric resource management across multiple devices on a large tabletop, and supported various gestures to configure activities and visualize available resources as well as pairing information. The Unadorned Desk [36] also augmented the physical space (on the desk) around the display as an input canvas for the desktop, while providing no feedback or on-screen feedback, rather than feedback on the desk. Touch input was widely used to indicate the user's focus on these devices, and such past work also suggests new ways that multiple devices can work together in a desk-centric environment.

Other approaches have used peripherals to transfer data between devices. Playing off traditional drag-and-drop techniques, Pick-and-Drop [83], Pipet [67], and Slurp [108] used handheld devices such as a pen to pick (i.e., copy) documents from one device and drop (i.e., paste) them to another. Stitching [39] improved this, using information from a stylus to automatically calibrate the spatial relationship between a pair of devices. In GroupTogether [66], the tilting of one's tablet towards another in close proximity transferred data between them, whereas with Throw and Tilt [18] and Touch + Interact [34], tilting and flicking were combined on a smartphone to transfer data. Other phone-based transfer techniques have included gesturing towards the display to be transferred to [9], tapping one's mobile phone on a display [88, 89], or taking a photo of one's workspace [10]. Still others have explored how the user *themselves* can take on the role of the transfer medium. Within LightSpace [103], a user can cup her hands to carry data between devices. In a similar vein, DynaWall [25] uses one's hand as the vessel to move content from one side of a large display to another. Turner et al. [98] took this further, using the eyes as the transfer mechanism. Gluey [92] displayed a 'clipboard' in a head-mounted display and enabled fluid content migration across devices by head movement. Although many techniques have used touch-input or an additional peripheral, e.g., DynaWall and LightSpace [25, 103], the present work explored how the *user* could be utilized as the transfer medium, supporting users as they looked, touched, or otherwise bridged devices.

2.2 Proxemics

The field of proxemics focuses on how the distances between users can mediate interpersonal interactions [32]. Within the context of HCI, the distance between a user and device [8, 66, 99–101], between multiple users of a single device [90], between multiple users and their personal devices [19, 45] and the movements, orientations, and identities between the users and devices [4, 26, 80] have been utilized to enrich interaction. One's proximity to a device has been used, for example, to modify the level of detail and content [21, 23, 44, 54, 63, 101], change the rendering and abstraction of information [1, 46, 58], or enable zooming within an interface [35, 54]. When device micro-mobility behaviors (i.e., orientation and tilt) were also utilized, a variety of collaborative cross-device sharing techniques were possible [66]. With proxemic-aware controls, the relative distance between smart devices and a user (who was holding a tablet) selectively revealed functionality [57]. The present work drew most inspiration from proxemic-aware controls, which were an examination of inter-device proxemic relationships (i.e., between the tablet and the smart device). Thus, this work focuses not only on the relations between users and devices, but also on relations between devices themselves. Previous work has focused on the cross-device interactions between multiple personal devices (e.g., smartphones) based on the distance and orientation [19, 45]. Illuminated by the explorations into the proxemics of devices [65, 66], the present work proposes that the proxemic relationships between multiple devices of different types can provide unique opportunities when considered within the context of a single user's device ecology.

Another focus within the research literature has been on the size, shape, and direction of the regions that devices occupy. Many projects have divided the interaction region into rectangular strips in front of a device, as they enable user interest to be inferred, or have appropriated the space immediately adjacent to a device (e.g., [13, 54, 99–101]). Others have proposed the use of user-focused circular or oval regions to facilitate interpersonal communications (e.g., [44, 66]). Within the present exploration, the notion of such zones is harnessed and extended such that zones differ not only in location or function, but also based on the type of device in question.

2.3 Paper and Digital Information Management

The need to efficiently manipulate and utilize multiple documents has become ever more prevalent with the continued push towards the paperless office and the ease of access to digital documentation and knowledge. Early work by Kirsh [53] highlighted users' tendencies to arrange documents in their environments such that they increase their success in conducting visual search tasks and decrease the number of choices and options they have available when seeking information. In a study of desk organization in professional office workers, Malone [61] found that environment organization facilitated the finding and reminding of information via piling. Worker's organizational strategies also harnessed the spatial closeness of information to the user and frequently used office appliances such as telephones. Sellen and Harper [91] noted the importance of piling, archiving, placing, shifting, and so on, when it came to manipulating multiple documents, as did Hong et al. [40], Takano et al. [95, 96], Mizrachi [71], and Bondarenko and Janssen [7]. Kidd [51] postulated that such behaviors enable one to demonstrate the progress they have made (i.e., decreasing pile size). The spatial organization of files also acts as a contextual cue and as a physical language that can be manipulated by the user. Work by Cole [17] focused on the types of information utilized throughout the day (i.e., personal work file, action information, and archive storage) and noted that archived information was often unorganized and spatially distant from the user. Furthermore, Sellen and Harper [91] defined notions of hot, warm, and cold documents, which demonstrated associations to the proximity of the user. The insights from this work underscore the importance of spatiality to cognitive processing and document management and the need to have relevant information close to the user. When designing techniques to facilitate interaction with multiple devices centered around a desk, as this work does, such projects underscore the importance of considering how spatial organization and information (i.e., device) proximity can facilitate new interaction techniques.

Others have focused on understanding how tasks are divided and distributed across multiple devices. In videotaped observational sessions, Oulasvirta and Sumari [77] found tasks were assigned based on device capabilities. Similar results were found by Karlson et al. [49] using device logging software. Grudin [30] interviewed users who had multiple monitors and found activities were distributed across monitors by task instead of by device capabilities. Through interviews, Dearman and Pierce [20] also found a primary/secondary device divide, but found functionality was continually assigned and revoked across one's device ecosystem. Interviews conducted by Santosa and Wigdor [86] identified two new interaction patterns centered on the physical properties of a form factor (i.e., the smartphone as a helper for quick activities such as calculation or search) and viewer/controller metaphors (i.e., the smartphone as a remote for a presentation or a music player). This prior work provided justification that the SMAC model would greatly benefit from utilizing the information that can be gleaned from the relationships between devices themselves.

The vision of multi-device desk-centric environments utilized within this exploration centers around harmonious interaction. Different methods of augmenting the desktop via projected surfaces or tabletops have been investigated by Kane et al. [48] and Bi et al. [6]. Although Myers [72] proposed the use of handheld devices as tools to support desktop computers, the community is still lacking knowledge about how user-and inter-device proxemic information could be used within a unified model to improve the multi-device workflow. Thus, we analyzed barriers that still exist in current workflows and summarized an interaction model for new and existing techniques in multi-device desk-centric environments.

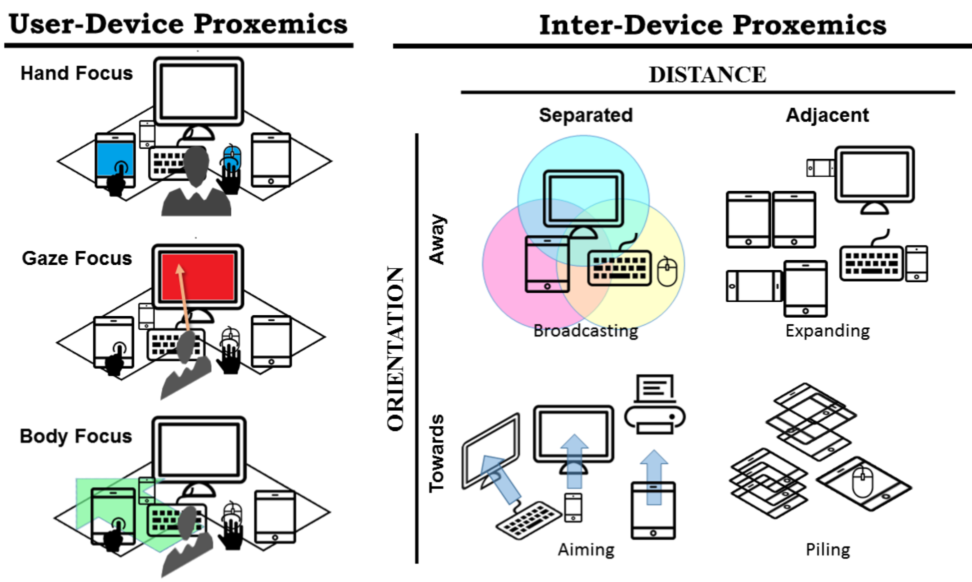

Fig. 1. SMAC: A Simplified Model of Attention and Capture mediated by user-device proxemics (i.e., hand, gaze, and body focus) coupled with inter-device proxemics (i.e., device distance and orientation). Using this model, a computing device would obtain user- and inter-device proxemic states and alter its behaviors based on this information. Details about how to build a system that follows SMAC principles are provided in Section 4. Icons ©www.easyicon.net used under CC BY-SA 3.0.

3 SMAC: A SIMPLIFIED MODEL OF ATTENTION AND CAPTURE

Inspired by existing work that identified the barriers that exist while using multiple devices in the same workflow [78, 86], we propose SMAC, a Simplified Model of Attention and Capture to assist in the designing of cross-device interactions in desk-centric settings (Fig. 1). SMAC is based on a combination of traditional user-device proxemic principles, i.e., hand, gaze, and body focus, coupled with the notion of inter-device proxemics, i.e., inter-device distance and orientation. It also makes use of postural proxemic information to infer user intent.

We thus propose sampling the multidimensional behaviors of a user with respect to the current device being touched, the current focus of the user's gaze, and the current orientation of the user's body, to better understand the targets of her past, current, and future behavior and attention. In addition, we propose the use of the rich information that can be gathered from the relationships between devices themselves, i.e., distance and orientation. SMAC uses this user and device-based information in concert to develop a unified understanding of the current state of each device in a multi-device environment, similar to the window manager in a traditional desktop environment. SMAC provides design guidelines of how such a system should capture and route a user's actions to various devices.

Although significant effort has been expended to develop interaction techniques for desk-centric environments [6, 47, 72], the community still lacks knowledge about how these systems could be unified into a single model to increase user understanding within a multi-device environment. SMAC thus contributes an interaction model that clarifies how multi-device desk-centric environments should respond in different scenarios and the benefits that users could attain when performing desktop computing tasks in such environments. It is our hope that SMAC will be used to design future DUI systems and create interaction techniques, as we have done and describe in Section 4. Further, the transitions between different cells in the interaction model should be able to serve as the common 'events' that users employ to delimit their interactions in multi-device environments, similar to how users today understand window and keyboard focus, clicks, double-clicks, and window size changes in WIMP environments. To this end, we provided not only the set of interaction states, but also a description of what changes in those states might mean within a multi-device system.

3.1 User-Device Proxemics

Users' behaviors with respect to one another, and to devices, form the basis of the application of proxemic principles to HCI. Traditionally, a user's presence or distance from a device was of interest, as in Vogel and Balakrishnan's [100] early work with wall displays. In this early work, the user-device relationship was mediated by four distance-based zones. When considering a desk-centric environment, however, such distinctions are limiting as distance varies little, but focus varies greatly. For example, a user may sit in front of a monitor with her body facing forward, type using the keyboard in front of her, and not change her distance to the monitor while she is gazing at a secondary document shown on a tablet to her left. In this scenario, the user would remain within an intimate distance to the monitor [31], but change her focus from one device (i.e., monitor) to another (i.e., tablet). In this case, the user-device proxemic relationship is mediated by focus, not distance.

Although some focus changes could be explicit (e.g., moving or tapping on a device), others might be implicit (e.g., turning towards another device), suggesting that it is important to detect and understand user actions that are not primarily aimed at interacting with the system [87]. We thus propose considering the granularity of focus that a user can have with various devices, i.e., the present target of the user's *hands*, the focus of the user's *gaze*, and the area in front of the user's *body*. We describe each of these facets next.

3.1.1 Hand Focus. The location of the hand, the objects the hand is touching, or the objects the hand is pointing to, have been long used to indicate what elements in the environment have the users' attention. Building on this, a device has *hand focus* when it is touched by the user's hand and is (i) being used for input (e.g., typing, writing, moving) or (ii) is in direct control by the user (e.g., being held for reading). Thus, hand focus indicates the user's direct intention to temporarily or continually operate a device or interact within information, similar to Yoon et al.'s view of intent with micro-mobility [106]. Because hand focus indicates input intent, when used in combination with other foci, such as gaze or body focus, hand focus should be considered the dominant form of focus.

In addition to input data generated by the hand being processed on the target device, there are instances where it may be useful to forward said input to other devices (e.g., the user types on a keyboard and the system forwards the characters to another monitor).

3.1.2 Gaze Focus. With gaze focus, the user's eyes or head are used as a cue to indicate what she is looking at. Therefore, a device has *gaze focus* when the user is looking at the device. In its simplest form, this has been used to, for example, switch among multiple monitors while typing on the single keyboard [22]. Gluey used gaze focus to obtain the status of the target device, e.g., the printing tasks of a printer [92]. Unlike hand focus, users often move their gaze focus away from, and then quickly back towards, the primary device or document they are working with, e.g., to copy information, search for a relevant document, and so on. As gaze focus is less intentional and informative for fine operations compared to hand focus (e.g., a user can get distracted and temporarily look away), we consider it as an indicator of the user's attention, rather than as an explicit input method to control a system.

Unlike prior work [22], we propose that gaze focus should serve only to specify the output target, but not necessarily the input target. Whenever a device receives gaze focus, it should become the major output device, thus the user should see feedback about her actions on the device. For example, if a user is typing an email on her primary monitor, and temporarily looks over to another screen to review a sales figure, the keyboard input should *not* be redirected to this second screen. However, it could be useful to show the user what she is typing on this second screen, even as it continues to be routed to the original input window. In this example, feedback is provided to the user about her change in gaze focus and each input action made to the device(s) with hand focus. Thus, changes in gaze focus are asymmetrical: they specify targets for *retrieving* or *displaying* information but are not sufficient to redirect the target of input streams.

3.1.3 Body Focus. With body focus, the shoulders and torso of the user are used as an indicator of attention. Devices within the area immediately in front of the user's shoulders thus have the user's *body focus*. As it takes longer to move one's torso than her hand or head, body focus provides a sustained, continual clue with regards to the current intent of the user. If, for example, the user suddenly picks up a tablet laying on a side table while keeping her body towards her primary computing device, this may indicate that (i) use of the tablet might be for a temporary purpose and the user may be returning it back to its original location soon or (ii) the actions invoked on the tablet are supporting the primary task occurring on the primary computing device. On the other hand, if the user moves the tablet in front of her body (or turns her body towards the tablet), it is likely that she is going to use the tablet for an extended period of time and may want to transfer the task and interactions from the primary computing device onto the tablet. Thus, body focus could be a strong indicator of the importance of that device to the current workflow. When a device gets body focus, for example, it could awaken and become fully operable; when it is outside of body focus, it could show a simplified UI to help the user perform common actions.

Body focus echoes past work on hot, warm, and cold zones [91, 97], in which the user's desk was divided into regions based on whether documents were within, at, or outside the user's reach. In the present work, we propose the use of a sector-shaped area determined by the orientation-based body focus to indicate the primary workspace, instead of using the arm reach. This allows the zones to change in real time based on a user's current physical arrangement.

Compared with hand or gaze focus, body focus has the least flexibility in providing fine input to a system and conveys less information when used without other foci. As a result, it is another indicator that, when used with gaze focus, may affect the interface of a device. In such cases, the content will still only be modifiable when the device receives hand focus.

3.1.4 Summary. When the combination of these three foci is fully-crossed, much information can be gained about the users' intent without explicitly asking them (Table 1). For example, a device which has the users' hand, gaze, and body focus, is clearly the user's focal device, similar to how a user points her head and body towards a desktop computer while engaged in a focused input task. Other combinations of foci enable for more interesting inferences. For example, if a device gains gaze focus, but does not gain body nor hand focus, this could indicate that the user is quickly consulting the device (e.g., checking the calendar events on the phone), while still expressing a desire to continue to use the previous working device. The state may also indicate the user's desire to bring information from the working device to the target device that is being looked at (gaze focus) or bring information back from the temporarily gaze-focused device to her current working device. We explore combinations of these three foci, along with the inferences they could allow, in more detail later.

Table 1. The eight possible states of user intent based on the current hand, gaze, and body focus of a user with relation to the devices she has in her workspace.

	Gaze Focus		No Gaze Focus	
	Body Focus	*No Body Focus*	*Body Focus*	*No Body Focus*
Hand Focus	Device content is most relevant to the task at hand, device has attention and device is accepting input from the user.	Device content is least relevant to the task at hand, but device has attention and is accepting input from the user.	Device content is likely relevant to the task at hand, device does not currently have attention, but device is accepting input from the user.	Device content is least relevant to the task at hand, device does not have attention, but is accepting input from the user.
No Hand Focus	Device content is relevant to the task at hand, has attention but device is not accepting input from the user.	Device content is least relevant to the task at hand, has attention but is not accepting input from the user.	Device content is less relevant to the task at hand, does not currently have attention and is not currently accepting input from the user.	Device content is least relevant to the task at hand, does not have attention and is not accepting input from the user.

3.2 Inter-Device Proxemics

Similar to traditional views of user-device relationships, the devices within multi-device environments can also be characterized by their spatial relationship, including proxemic distance and orientation (Fig. 1). In past work, Marquardt et al. [63] proposed five inter-device proxemic dimensions: location, distance, movement, orientation, and identity. These dimensions supported their investigation into the sequential stages of a user walking towards a semi-public wall display with a handheld device, and collaborations among users with personal devices. However, in their work, the user's personal device served as a proxy for the user themselves; the distance between multiple personal devices was taken to be the distance between users, and the distance from a personal device to a wall display was taken to be the distance of the personal device's user from that display. Thus, a 1:1 relationship between each user and personal device was assumed. In the environment that formed the basis of our explorations, there is instead a one-to-many relationship between the user and devices. We thus propose an expansion of the notion of inter-device proxemics, in which the spatial relationship of devices is a function of the user's placement of them, rather than an indicator of the position of the user herself.

3.2.1 Distance. We designate devices' relative distance by the presence or absence of physical contact between them (*adjacent* or *separated*). Marquardt et al. [63] suggested three stages for inter-device interactions in larger settings, including *"awareness of device presence, reveal of exchangeable content, and transfer methods"*. Considering the reduced physical space at a desk, we have condensed the first two notions. Physically touching devices together (*adjacent*) is a strong indicator that the user wishes to connect two devices and share information, thus a system should respond to this action by displaying a sharing interface and feedback about the connection. Alternatively, devices could be *separated* by being placed away from each other, isolated in space. This is the natural and default state of most devices around a user. Such devices should still be aware of each other's presence and location but require the user to explicitly confirm the pairing to avoid unintended input.

Table 2. The four states that a device could be in when considered along the dimensions of distance (separated, adjacent) and orientation (away, towards).

		Distance	
		Separated	*Adjacent*
Orientation	*Away*	**Broadcasting**: The lack of directional and proximal relationships indicates that information should be intermittently transmitted between devices (if at all).	**Expanding**: The lack of a directional relationship but close proximity between devices indicates that they are grouped and should continually share information.
	Towards	**Aiming**: The directional relationship between the devices indicates that information should be continually transmitted from one device to another.	**Piling**: The directional relationship and close proximity between devices indicates that they are grouped, and their information should not only be continually shared, but also tightly coupled.

3.2.2 Orientation. The relative orientation of devices also reveals the user's intention to pair or group them, similar to how people orient themselves towards another while conversing [66]. For example, users control a TV by aiming a remote controller *towards* the TV, in accordance with the direction of the lettering or symbols on the buttons. The same can be said for keyboards (i.e., the direction of the lettering or symbols on the keys indicates the 'front' of the device). Mobile devices, including phones and tablets, would use the 'back' camera to dictate the direction of the device. We consider a device to be oriented *away* from another if it is not aimed *towards* it. For example, two tablets resting next to each other face-up on the same desk (with both back cameras aiming at the desk) are considered to be aimed away from one another. Thus, the relative proportion of device orientation space occupied by *away* is much larger than that of *towards*.

When the user orientates a handheld device *towards* another device, the devices may become linked, so that the in-hand device can control the remote device to which it is being pointed. When a controlling device is turned *away* from another device, both devices should disconnect. It should be noted that user intent can be difficult to infer in this situation, as devices might be placed at an ambiguous orientation (e.g., 'aiming' at a 45-degree angle). To assist with this, the system should consider additional context information, including the application history, and the dynamics of the distance and intersection angle of the user's reorientation of the device action.

3.2.3 Distance and Orientation. By coupling these two dimensions, four inter-device proxemic states emerge: *broadcasting*, *expanding*, *aiming*, and *piling* (Table 2). We present the rules to determine each state and suggested responses or feedback for devices in each state.

Broadcasting. When devices are *separated*, they are likely too far from each other to be used at the same time for the current task. When they are turned *away* from each other, it is likely that the user does not intend for one to exert control over another. The combination of separated and away creates the broadcasting state. This state is the default state for all devices in an environment until they point towards each other or come in closer proximity to each other. Broadcasting devices could provide feedback about nearby input devices or possible displays for output.

Prior work has explored additional ways to utilize this state. For example, Marquardt et al. [63] explored the usage of dynamic notifications about the presence and location of nearby devices and designed continuous proxemics-dependent stages for seamless transitions. Relate Gateways [26] used a compass metaphor to provide mobile devices with a view of the services available on the edges of the display.

Expanding. When devices are *adjacent* but facing *away* from each other, they can be said to be in a state where they are close enough that they are likely being used for the same task, but neither one is more important or in control of the other. In essence, the devices have been grouped via distance. This could be thought of as a modern-day equivalent of placing thematically-related loose-leaf sheets next to each other while writing a report. When in such a state, each device should provide the user with feedback about the devices which are nearby or the information that such devices could share.

Placing devices side-by-side on the desk, for example, could mean that the devices should be linked together, either for a larger, continuous workspace or to make information communally available, as it may suggest a user's desire to compare or share information across devices, and use them to complete the same task together. This positioning could create a continuous space for a single or multiple users, as demonstrated by Hinckley et al. [38], Lucero et al. [60] and Chen et al. [11]. It also reflects the potential desire for data to be transferred across the edges of the devices [39, 66].

Aiming. When devices are *separated*, yet one device is orientated *towards* another, they are said to be in a state where one device is controlling or providing input to another. Inspired by how we use traditional TV remote controllers, one device becomes the source device (e.g., remote controller) and is used to control the other target device (e.g., TV). In such a state, users may provide input to the source device and anticipate feedback on the target device [5, 57], or wish to operate directly with the source device after it is connected to the target device [8, 73]. This state is especially useful when a single input device is mapped to multiple output devices.

Others such as Chong and Gellersen [15] explored the use of this state to associate devices to each other, i.e., pointing by approximation (e.g., an infrared beam from the TV remote controller) or pointing with precision (e.g., a laser light that supports selection).

Piling. If devices are *adjacent* and facing *towards* each other, they are most likely being piled or stacked. Similar to the piling that can occur with physical paper, this can be an implicit cue that all of the information or features that are part of the pile contents are *highly* related. With digital devices, the distance (*adjacent*) between the devices is a strong cue that a connection needs to be established between devices, and the orientation (*towards*) suggests that one device is exerting control over the others and thus functionality and information should be passed to, and available on, the topmost device.

Like paper, devices are usually piled in the same direction (i.e., all face up or all face down). However, because other postures such as devices piled face to face or back to back are also possible, the user may want to group the devices within the pile by doing so. When users manually flip the device on top or rotate the device to be perpendicular to another device in the pile, this could be used as a further indicator of importance or intent.

3.3 Summary

As illustrated by SMAC (Fig. 1), when the factors of user- and inter-device proxemics are coupled, they can be used to implicitly deduce user intent and thus shape interaction and feedback. This interaction model allows developers and designers to utilize the state of the environment and the user by harnessing the natural behavior that the user exhibits over, and exemplifies with, her devices. The choice of user-device proxemics, i.e., hand, gaze, and body focus, reflect a continuous design space from the most explicit touch input to implicit body orientation. Other foci including head orientation and foot focus could be added to this model in the future, as they become more commonplace ways to interact with devices. As for inter-device proxemics, i.e., the distance and orientation between devices, because the area on many desks is restricted, SMAC utilizes the two most appropriate dimensions for this area. Additional granularity within these dimensions could be

Table 3. Summary of how existing interaction techniques from the literature fit into SMAC.

	Orientation: Away		**Orientation: Towards**	
	Distance: Separated (*Broadcasting*)	Distance: Adjacent (*Expanding*)	Distance: Separated (*Aiming*)	Distance: Adjacent (*Piling*)
Focus: Hand	*Conductor* [33]; *Gradual Engagement* [63]; *Orienteer* [19]; *Pebbles* [72]; *Relate Gateways* [26]; *SyncTap* [84]; *Tracko* [45]; *United Slates* [11]; *Weave* [14]; *XD-Browser* [75].	*Conductor* [33]; *HuddleLamp* [80]; *JuxtaPinch* [76]; *Panelrama* [105]; *Pass-Them-Around* [60]; *Stitching* [39]; *United Slates* [11].	*Code Space* [9]; *CTAT* [21]; *GroupTogether* [66]; *Touch Projector* [8]; *PhoneTouch* [88]; *Pipet* [67]; *Proxemic-Aware Controls* [57]; *Proxemic Interaction* [4]; *Slurp* [108]; *Throw and Tilt* [18].	*PaperTab* [97]
Focus: Gaze	*LookPoint* [22]; *MAGIC Pointing* [107]; *Pupil* [50]; *ViewPointer* [93]; *Gluey* [92]; *Gaze Positioning* [98].			
Focus: Body	*Chair Interaction* [79]			

explored if this model is extended to apply to larger or smaller spaces than the size of the average desk. For multi-user scenarios, SMAC could be extended to include inter-user dimensions such as F-formations [66]. Although the present exploration focused on desktop-based work, SMAC can be extended to other situations as well. For example, in co-working or public spaces, as long as a user is able to identify one device as her 'primary' device, the device and user focused facets of SMAC would still be applicable.

As illustrated in Table 3, much of the existing work on multi-device interaction techniques naturally fits within SMAC. Compared to prior projects that explored the broadcasting, expanding, and aiming states, the piling state has the greatest potential to be explored further. In addition, since most of prior projects used touch input and in-air gestures, other projects using gaze focus [22, 50, 93, 107] or body focus [79] could also be enhanced by the use of the different inter-device states. The gaps left by prior work afford many new possibilities in terms of multi-device interaction techniques. To further explore the newfound opportunities that focus on user- and inter-device proxemics may have on interaction, next, we detail the design and implementation of the prototype system, OmniDesk. The prototype system explores interaction techniques when the three user-device foci (i.e., hand, gaze, and body focus) and four inter-device states (i.e., broadcasting, expanding, aiming and piling state) are harnessed.

4 EXPLORING USER- AND INTER-DEVICE PROXEMICS

The main contribution of the SMAC model lies within its ability to unify and extend existing distributed user interface interaction techniques. To better understand the interaction opportunities

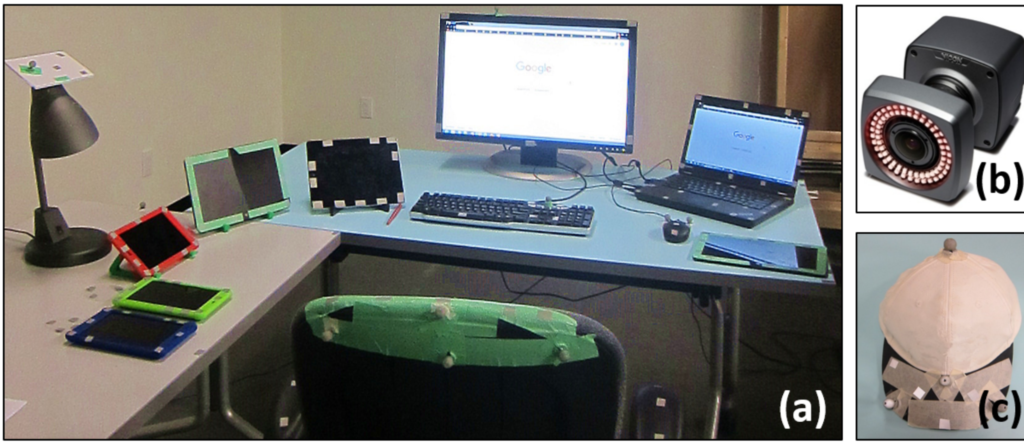

Fig. 2. Hardware configuration for our prototype system, OmniDesk. (a) Overall view of the environment, including an L-shaped desk, an office chair, a lamp, tablets, laptop/desktop computers, and a keyboard and mouse. (b) Six Vicon Bonita motion capture cameras mounted to the ceiling were used to track the user and the devices. (c) The baseball cap with retro-reflective markers tracked the user's head movement.

afforded when considering user-device proxemic states (i.e., hand, gaze, and body focus) and inter-device proxemic states (i.e., broadcasting, expanding, aiming, and piling state), herein we explore a collection of techniques from the literature, along with novel variants of them that make use of the state information SMAC provides. For example, the Adjustable Reader and Device Radar demonstrate how "United Slates" [11] and "Conductor" [33] could be enhanced by using additional user- and inter-device proxemic states. Other techniques such as the Gaze Search and Piling Hub reveal the benefits of combining multiple states or transitioning between states.

4.1 OmniDesk System

To realize these techniques, a prototype system, *OmniDesk*, was built (Fig. 2). It used six Vicon Bonita cameras situated around a desk-based environment. The environment contained a laptop, a secondary monitor, six tablets, a lamp, a mouse, and a keyboard. Each device in the environment was 'tagged' with retroreflective markers to allow the Vicon system to triangulate the position and orientation of each device (i.e., inter-device proxemics). As for the piling state, since the topmost tablet usually blocks the markers of other devices underneath it, we placed additional markers around the tablets and tracked the movement and presence or absence of the markers of every tablet to determine every tablet's final location. OmniDesk considers hand focus to be explicit input on a touch screen, the keyboard keys, or the mouse. More fine-grained sensing of the hands and fingers could be used in the future to enable additional interactions. To determine where the user was gazing, a baseball cap was outfitted with retroreflective markers. Although the use of markers on a baseball cap would not allow for the detection of a change in the movement of a user's eyes, as the current focus was gaze switching between devices, rather than the gaze movement within a device, this abstraction was deemed acceptable. Pervasive eye-tracking solutions [50, 93], however, could be applied in future work and provide more robust sensing results. To determine where the user's body was facing, a computer chair was retrofitted with a variety of retro-reflective markers. When the user turned around to different devices in the chair, said markers allowed convenient detection of the current body focus.

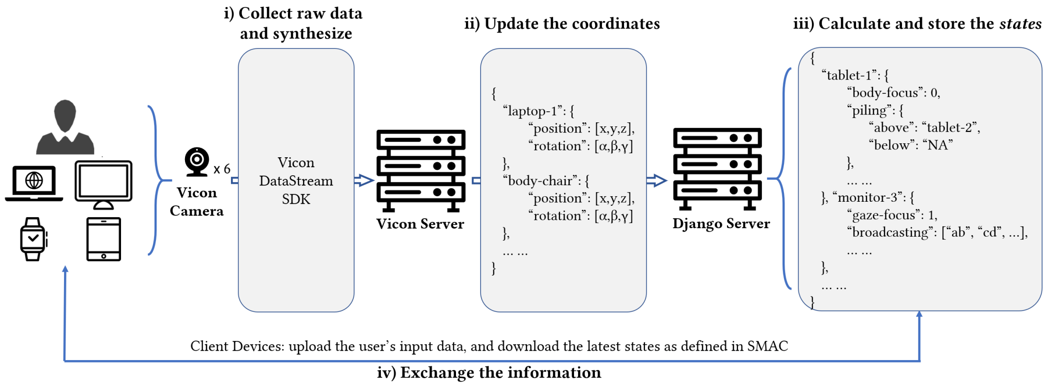

Fig. 3. The software infrastructure of OmniDesk. i) The raw data is collected and sent to the Vicon server via the Vicon DataStream SDK. ii) The Vicon server updates the coordinate of every object, including the positions and rotations (Euler angles), and sends it to the Django server. iii) The Django server calculates the user- and inter-device states of each client device. iv) Each device can retrieve the updated status information and upload any user input data by periodically communicating with the Django server.

There are three major components in the software infrastructure: a Vicon server, a Django server, and multiple client devices (Fig. 3). The Vicon server collects the raw positional data of each digital and analog devices using the Vicon DataStream SDK and updates the processed position and rotation information to the Django server via HTTP POST requests. The Django server then calculates the user- and inter-device state information as defined by SMAC. Each client device uploads local changes (e.g., the user's input on the touchscreen or keyboard) directly to the Django server and retrieves the latest user- and inter-device proxemic information from the server via HTTP POST requests. A small subset of simulated web-applications was developed as proof-of-concept applications, including simplified versions of Microsoft Word, Excel, and PowerPoint, as well as a PDF reader. As web-based toolkits can provide unified framework solutions across platforms and devices [24, 62, 75], OmniDesk used a simplified version of such architectures for the web-applications.

4.2 Broadcasting

When the devices within one's digital ecosystem are physically separated and pointed away from each other, many new interactions are possible. In the four examples that follow, we explored how this single state of inter-device information could be combined with hand-, gaze-, or body-focus to enable new workflows and interactions.

4.2.1 Device Radar. As each device within a multi-device environment is typically isolated and unaware of the existence or proximity of other devices, systems that have an awareness of the capabilities of other devices can utilize inter-device distance to visualize said information [65]. Much like how the 'hover' state is used to preview what will happen if a user 'clicks' with a mouse, device closeness and separation can provide previews of the functionality supported by other devices.

With the Device Radar, when a device is broadcasting, it signifies the possibility and result of increasing the proxemic coupling between devices and the user. For example, if a user receives an email on her tablet that has a limited screen space, the tablet can display a monitor icon on the tablet's interface to alert the user that a more suitable device is available to compose her

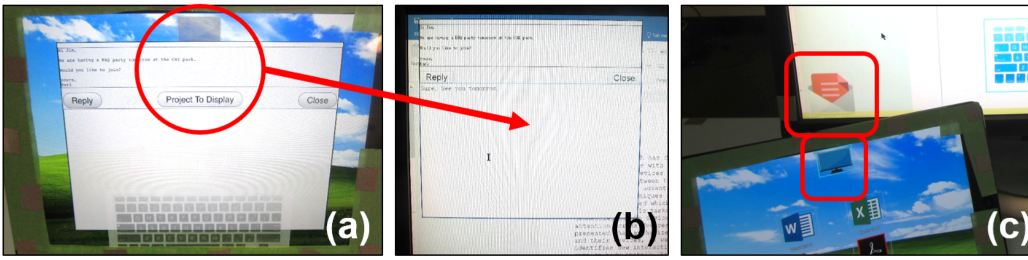

Fig. 4. The *broadcasting* state and *hand* focus are used to enhance the experience of receiving and viewing emails on multiple devices with the Device Radar. (a) The monitor icon (top) and the keyboard icon (bottom) on the tablet indicates that these devices are ready for connection. (b) By tapping the "Project to Display" button, current and future notifications on the tablet will be cast to the monitor. (c) The location of the notification on the monitor corresponds to the physical location of the tablet.

email response (Fig. 4a). Because all OmniDesk devices continually broadcast their capabilities and locations, the user can explicitly set up a connection between the monitor and tablet by pressing the "Project to Display" button. Thus, all future emails that the tablet receives will appear on this monitor in the form of a toast dialog that is located parallel to tablet's physical location (Fig. 4c).

This example interaction is a fusion of Conductor [33] (i.e., similar devices placed in the same location provided the user with the opportunity to increase coupling via the on-screen user interface), United Slates [11] and Relate Gateways [26] (i.e., devices show dynamic icons along the edge of a display to indicate the presence and location of other devices in close proximity). The Device Radar harnesses the inter-device information attained from the broadcasting state to provide environmental device awareness and utilizes the user-device hand focus state to create a connection between the two devices, much like Conductor. It goes one step further, however, in that it visualizes the relationship between the tablet and monitor, similar to United Slates and Relate Gateways, to provide the user with persistent feedback about the connection.

4.2.2 Gaze Copy and Gaze Search. As users are quite comfortable using their eyes to indicate a target or area for their actions, instead of using their eyes to point [22], such information can be combined with the information from the inter-device broadcasting state to create natural selection and copying behaviors. In addition, hand focus has also been found helpful to trigger or confirm the gaze-input action [92, 98].

In the OmniDesk system, if a user wishes to copy content from one device to another, she can use a combination of gaze and hand focus actions (i.e., Gaze Copy; Fig. 5). With this procedure, the user first clicks or taps on the title bar of the application (hand focus; Fig. 5a). She then selects the target device by gazing at it (gaze focus; Fig. 5b) while receiving real-time visual feedback and finally confirms her intention by releasing the mouse or her finger (hand focus; Fig. 5c). With the Gaze Copy, users themselves can act as a clipboard, internalizing the movement of content across devices via their gaze and externalizing the selection and pasting of said content via hand focus. By harnessing our natural desire to use our eyes to look at targets and output devices, the user can create copies on multiple devices without needing to physically drag content across devices.

In addition, the user can perform gaze-directed search by selecting a keyword on the current device using the mouse or a finger (i.e., Gaze Search; Fig. 5d). To not interfere with the current content on the primary device being used, the user can gaze towards any other device in her environment to invoke a search. Search results from online websites, as well as the devices that are broadcasting in the workspace, are displayed in a split view on this target device. To enable the user

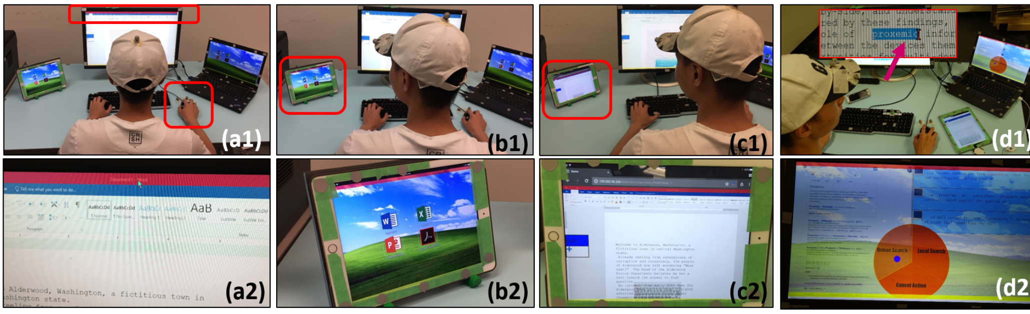

Fig. 5. Illustration of the *broadcasting* state and *gaze* focus. With the Gaze Copy, the user long clicks on the title bar of the original application window (a1/a2). The user then looks at the target tablet and the red indicator shows that this device currently has gaze focus (b1/b2). As the user releases the mouse, the application is copied to the target device (c1/c2). With the Gaze Search, the user selects a keyword with the mouse and looks at the target monitor to perform a gaze-directed search (d1). Results from online websites (left) and the local workspace (right) are presented using a split view (d2).

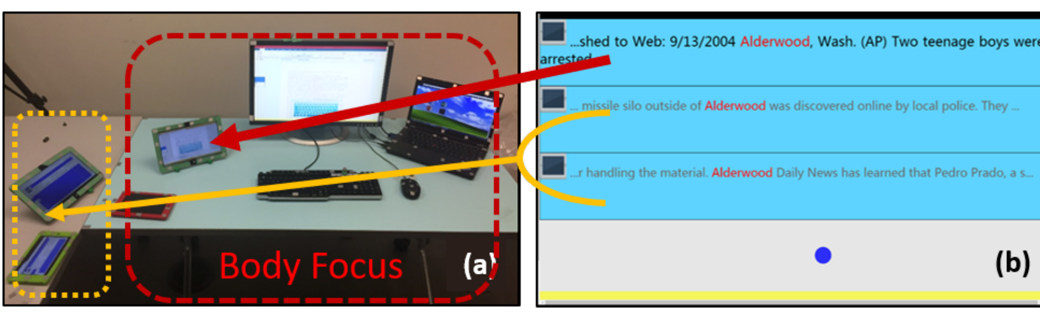

Fig. 6. Illustration of the *broadcasting* state and *gaze* and *body* focus with the Gaze Search. (a) There are two tablets in front of the user, which have body focus (red box), and the other two tablets do not have body focus (yellow box). (b) In the searching result, the found document that has body focus (red arrow) is displayed on top with different styles to help the user distinguish between them.

to browse through the search results, OmniDesk automatically redirects the mouse (or finger) input to this target device (hand focus). Because the search results come from the aggregation of many device-specific searches, when the user selects a search result that came from another device in the workspace, the corresponding device will blink to assist the user in finding it on her desk. This feedback is helpful when the device may not be within reach or when the device may be partially hidden under papers or other devices. With the Gaze Search, the user can internalize the selection of the device to search on by looking at it and externalizing the specification of the search terms via hand focus.

The Gaze Search can be further extended by integrating gaze and body focus information. For example, if documents or devices are within one's body focus but separated from the primary device, OmniDesk can infer that these documents or devices are relevant to the current task. Thus, search results that come from devices with body focus could be displayed higher in the result list (Fig. 6). This is an example of combining the three user-device foci: the user first selects the search terms (hand focus), then invokes the search on the target display (gaze focus) where the results are organized by body focus, and finally browses and acts on the results (hand focus).

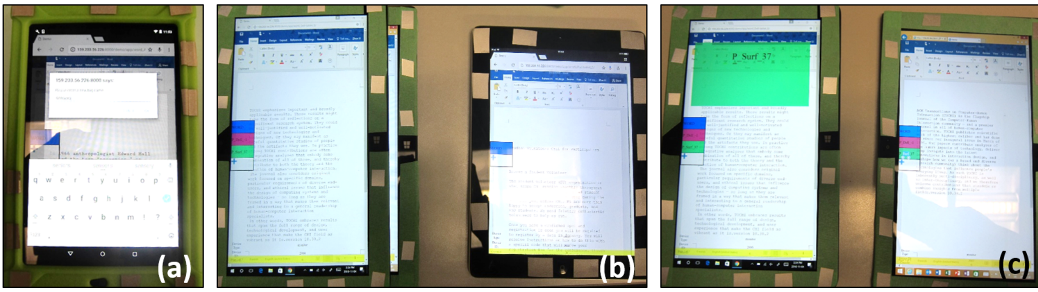

Fig. 7. Illustration of the *broadcasting* state with the Digital Labels. (a) A label is created manually. (b) Labels are automatically created for tablets in the same pile. (c) By tapping on one label, other devices with the same label blink.

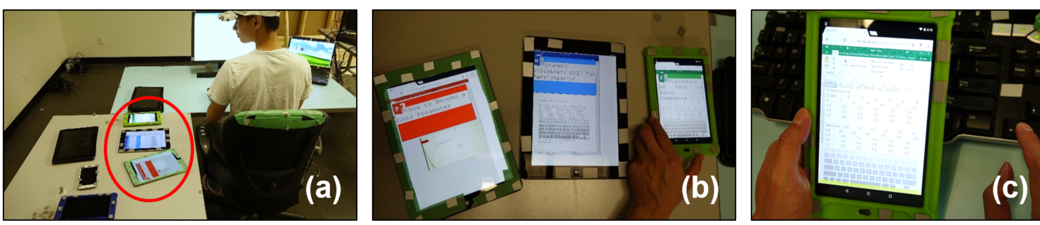

Fig. 8. The Summary View technique combines *body* focus with the *broadcasting* state. (a) Devices without body focus display the summary view, which includes a large title and panels that use color to illustrate the last application that was used. (b) When the user grabs a tablet and (c) places it in front of his body, this device has body focus and displays the original content, unchanged.

4.2.3 Digital Labels. Chen et al. implemented a "tagging" system, where tablets in the same pile could be automatically tagged to make use of the user's spatial memory to organize documents [11]. Inspired by Chen et al.'s work, along with how people create tags and categorize documents using physical sticky notes, we implemented the Digital Labels (Fig. 7). With these labels, the user can manually create a label on one device and tap on other devices to assign this new label. Labels can also be automatically generated for documents within the same pile (Fig. 7b). When a pile of documents is distributed on the desk, the user can find documents that were once in the same pile by tapping on one label - all nearby devices with the same label will start blinking (Fig. 7c). Similarly, by broadcasting recent changes to nearby devices, different copies of the same document can stay synchronized. Thanks to the broadcasting state and hand focus, this technique provides better connectivity and supports convenient navigation compared to traditional printed documents.

4.2.4 Summary View. As we naturally orient our bodies towards information and stimuli that have our attention, the Summary View technique utilizes the direction of one's torso (body focus), in combination with the broadcasting state of devices, to allow users to visually scan the devices in their environment. This usage of body focus was inspired by Schmidt's [87] vision of Implicit HCI - the OmniDesk system uses body focus as an indicator of the devices that the user may soon interact with and modulates the output on such devices accordingly.

Within the Summary View technique, the information displayed on a device's screen changes based on the current state of the user's body focus. Devices within one's body focus display the last content that the user interacted with, in an unaltered view. Devices outside one's body focus, however, display a summary of the on-screen content, using a larger title size and colored panels

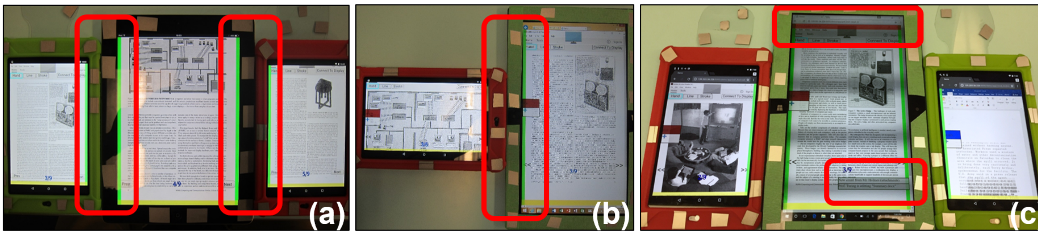

Fig. 9. Illustrations of the *expanding* state with the Adjustable Reader and Enhanced Adjustable Reader. (a) Adjustable Reader: Placing the tablet adjacent to another creates a continuous space. (b) Devices with different display orientations (portrait or landscape) are assigned with different types of content. (c) Enhanced Adjustable Reader with *gaze* focus: the red indicator appears when the device receives gaze focus, causing the notifications for this group of devices to be aggregated and displayed here.

to illustrate the last application that was in use (Fig. 8a). With these visualizations, the user can scan her environment to find a needed document by glancing at the summary views. If a desired document is found, the user can pick up the device or turn her body towards it. Picking up the device will display the entire document or application (unaltered view) and enable the device to receive input (hand focus). Turning one's body towards the device will show the unaltered view because the device has body focus but will not enable input. To avoid interruption, this body-based mode switching technique can be canceled, by tapping on the screen to temporally hide the summary view.

4.3 Expanding

When devices are located close to each other, yet are oriented away from each other, this is a reasonably clear indication that the devices should or do share some link, whether thematically in terms of content or functionally in terms of device characteristics. Next, we examine four techniques that harness the information that nearby, grouped devices afford.

4.3.1 Adjustable Reader and Enhanced Adjustable Reader. As it is common for users to spread paper documents on their desks to compare information across pages, the Adjustable Reader technique bootstraps such behavior. With the Adjustable Reader, whenever the user moves devices close to each other such that they are physically touching along one side, the currently viewed document will be visualized across all of the devices that are touching (Fig. 9a), similar to Chen et al. [11]'s "multi-page view" for tablets. In addition, the physical closeness of tablets can assign different types of content based on a device's status. For example, the user can view tables and figures on a landscape tablet and text on adjacent portrait tablets (Fig. 9b), while all three tablets remain synchronized and display the same article. A green sidebar is displayed on each 'expanded' device to indicate that an expanding connection has been formed with the other devices. Similar feedback features have been previously explored in prior work [11, 33, 105].

Although the inter-device expanding state has been explored in the literature, it has not made use of any of the user-device states that are equally important to consider. The fusing of gaze and body focus could extend this technique even further. With the Enhanced Adjustable Reader, when a user is working with devices that are in the expanding state, the notifications generated by each of these devices can be aggregated and propagated to the device that currently has gaze focus (Fig. 9c). This would be useful when using larger connected devices, or an array of smaller connected devices; in both situations, it is likely that the large 'display' created by joining all these devices may result in some content falling out of view, and the Enhanced Adjustable Reader could make

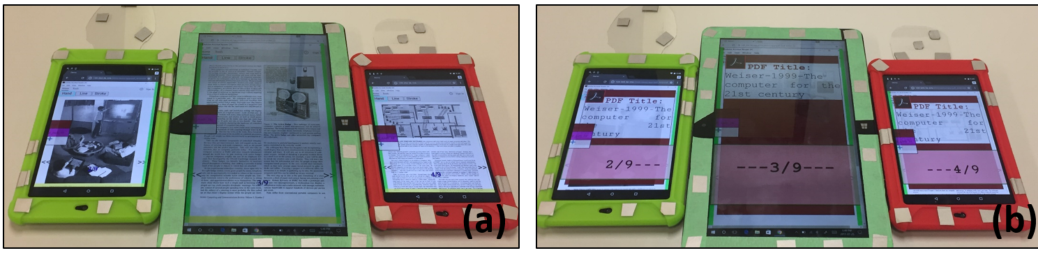

Fig. 10. Illustration of the *expanding* state and *body* focus with the Enhanced Adjustable Reader. (a) The devices present regular views if they have body focus and are ready for further interaction in the expanding state. (b) The devices present a summary view without body focus, which reveals their current expanding state.

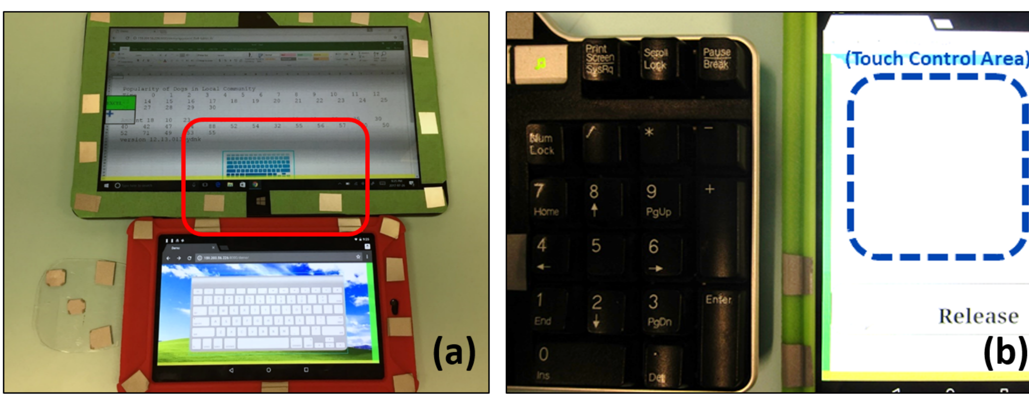

Fig. 11. Illustrations of the *expanding* state with: (a) the Expanded Keyboard, which creates a soft keyboard on the bottom tablet and displays a connection indicator on the top tablet, and (b) the Expanded Trackpad, which creates a touchable 'trackpad' area on the tablet to receive input from the user.

the notifications always visible. Further, for those devices that fall outside the user's body focus, the Enhanced Adjustable Reader could display summary information such as the application name or file when the user glances at them (Fig. 10b).

4.3.2 Expanded Keyboard and Expanded Trackpad. While the Enhanced Adjustable Reader harnessed the inter-device expanding state, along with the user-device states of gaze, and later body focus, it neglected to explore how hand focus could provide unique input possibilities. In what follows, we describe two uses of the expanding state, i.e., the Expanded Keyboard and the Expanded Trackpad, wherein different functionality is assigned to different devices such that one device supports another, i.e., transferring 'action' instead of information. Both techniques were inspired by HuddleLamp's assignment of different functionality to different devices when devices were grouped in a "huddle" [80].

With the Expanded Keyboard, when two devices are placed adjacently, with one above the other and within one's body focus, a soft keyboard will be automatically shown on the bottom device, thereby creating an ad-hoc 'laptop' (Fig. 11a). This not only allows the keyboard to be located closer to the user to increase her typing comfort, but also reduces occlusion issues that are often found when typing and viewing content on a single, smaller device. To ensure the user remembers that

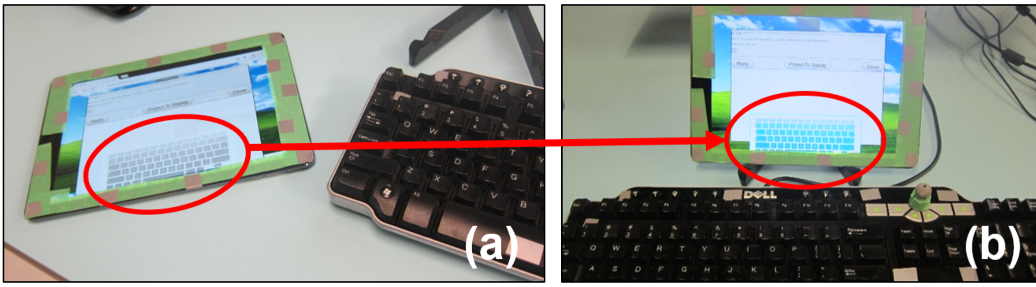

Fig. 12. The Universal Keyboard technique illustrates the transition from the *broadcasting* state to the *aiming* state. (a) The highlighted keyboard icon suggests an available connection to a keyboard nearby. (b) Placing the tablet behind the keyboard sets up the aiming state, and the icon will be highlighted. The keyboard icon has been enlarged for visibility.

the devices are connected, a blue keyboard icon will appear at the bottom of the topmost device. This technique is enabled by the expanding state, since both devices are exchanging information about their proximity and become 'one' device, and it is also supported by the user's touch gestures (hand focus).

As a variant of the Expanded Keyboard technique, if a keyboard and target display have already been connected, the user can activate an 'Expanded Trackpad' by placing another device adjacent to the keyboard (Fig. 11b). The target display will provide the user with feedback that the keyboard has been coupled to the additional device. The user can then click, drag, or perform gestures on the 'trackpad' (hand focus), and these actions will be synchronized to the target display, thus sharing information with the keyboard (expanding state).

4.4 Aiming

There are a number of new opportunities to consider when a multi-device environment has an awareness that devices are physically separated yet oriented towards each other. In addition to implementing a remote controller, i.e., a technique that allows the user to operate a tablet like a literal remote controller [8, 26, 57], we also explored how the use of additional information such as gaze focus to extend this classical concept much further. Next, we present three techniques that make use of the rich inter-device and user-device information that is available in multi-device environments.

4.4.1 Universal Keyboard. Within multi-device environments, it is common to find multiple tablets or external monitors, but only one keyboard is typically attached to a desktop computer [86]. In modern environments, a keyboard can be switched between devices using Bluetooth pairing, however, this is a long and tedious process. Past research-based systems have also described using gaze input or physical keys for keyboard/screen pairing [22, 92]. As the aiming state in SMAC enables for control to be exerted by one device to another, the Universal Keyboard supports free keyboard/screen pairing while providing visual cues of the connection.

If a user receives a new email on her tablet but the on-screen soft keyboard is quite small, she may prefer to use a nearby physical keyboard to compose a response. Because the keyboard is currently pointed away from the tablet and the devices are not adjacent, the keyboard is in the *broadcasting* state with respect to the tablet. The user will, as a result, see a keyboard icon on the bottom of the tablet (Fig. 12a). This icon signifies that she can aim the keyboard towards the tablet to provide input to the tablet. Once aimed, the keyboard icon will change to indicate that the user

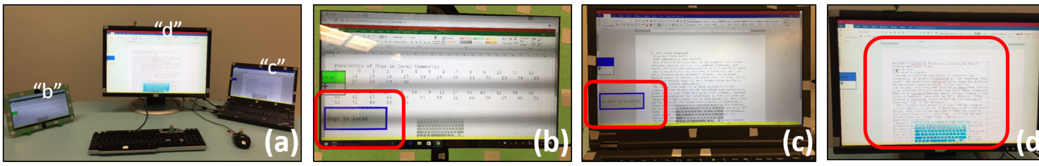

Fig. 13. Exploring the *aiming* state and *hand* and *gaze* focus with the Following Window. (a) The keyboard is aimed at monitor "d" and the user is typing on it. (b) If the user looks at tablet "b", the Following Window keeps track of her typed words on the keyboard. (c) Similarly, if the user looks at laptop "c", the Following Window keeps track of her typed words on the keyboard. (d) On monitor "d", all characters are processed normally so the system will send a copy to the Following Windows on the other devices based on the user's current gaze focus.

is aiming the keyboard towards the tablet and the output from the keyboard will be propagated to the tablet (Fig. 12b). This technique demonstrates how transitioning to the aiming state can enable fluid interaction with devices. As this technique relies on the aiming state and not only on the user's focus, it can also avoid some of the issues caused by unintentional gaze switching.

4.4.2 Following Window. While working on a report or paper, it is common to have a variety of documents, webpages, or applications open on different devices - with one or more documents being the 'main' document and others being used for reference or calculations. As one often continues typing in her main document, it is common to make transcription errors or typos. By combining the aiming state with hand and gaze focus, a new technique, Following Window, can alleviate this challenge.

With the Following Window, the aiming state and gaze and hand focus are used to create a temporary feedback window on the device that has gaze focus but is currently not in the aiming state. For example, consider the situation where the user is typing an article eyes-free using a Universal Keyboard (hand focus and aiming state) while she is looking at the reference document on another monitor. With the OmniDesk system, a preview window, i.e., Following Window, will appear on the output device that she is gazing at. Because the keyboard is aimed at the primary output device and not aimed at the secondary device she is looking at, and the keyboard has hand focus, the input will continue to be directed to the primary output device and will not be redirected to the secondary device. The secondary output device will display a small preview window that will show the last characters typed by the user (Fig. 13b). If the user were to maintain her gaze on the secondary monitor, aim the Universal Keyboard towards the secondary monitor and continue typing, the Following Window would disappear. With this technique, the user's keyboard input (i.e., hand focus) is processed based on the device with the aiming state, and feedback is displayed in a 'floating' window that follows the user's gaze focus to lessen the chances of typos.

4.4.3 SyncNotes. When annotations and notes are made on the physical documents, they are isolated from the user's digital world and thus hard to retrieve. In a digital workspace, where all physical documents are replaced by tablets, the user can aim one computing device at another to set up a projector-style connection. By using many tablets together with a limited number of external monitors, the user can rearrange her external monitor to match different tablets, with the connection signified by the proxemics between devices. This technique can also be expanded to send the content of the tablet to different location, such as a projector or a wall display.

To illustrate this concept, we designed a note-taking application named "SyncNotes". If tablets are brought into the center of the user's working zone (body focus), OmniDesk assumes that the user is reading the content it in an immersive manner [40]. Highlighted paragraphs thus become

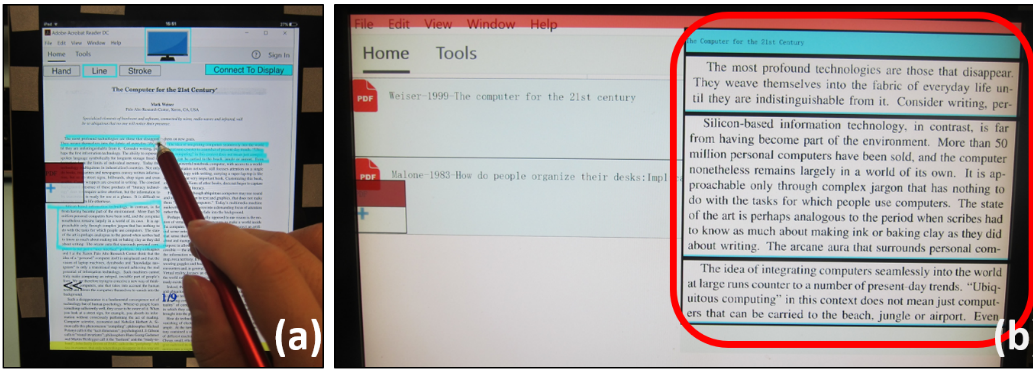

Fig. 14. Illustration of the *aiming* state and *hand* and *body* focus with the SyncNotes. (a) A user reading on a tablet and making annotations (three edit modes: hand, line, and stroke). After aiming the tablet at the target monitor, the highlighted icon on top indicates that a connection was made between the devices. (b) After returning her attention to the monitor, the user can see a list of the paragraphs she annotated on the tablet (highlighted by the red box).

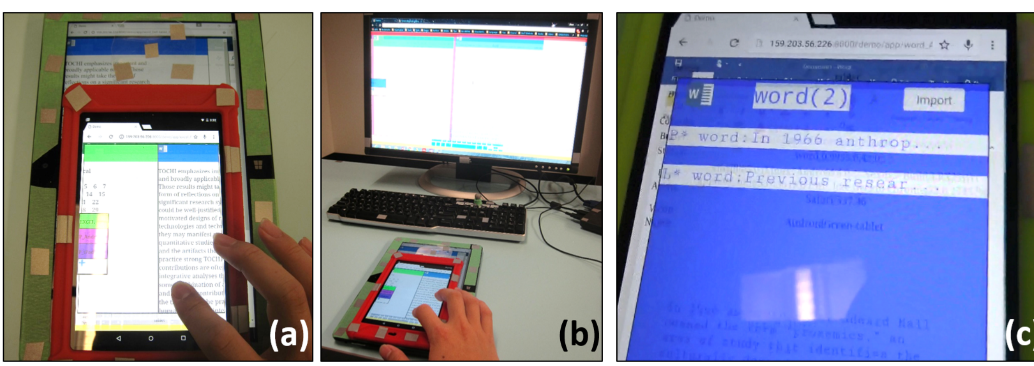

Fig. 15. Illustration of the *piling* state and *hand*, *gaze*, and *body* focus with the Piling Portal. (a) The user navigating a pile of documents on the topmost tablet. (b) The user projecting the document navigation UI on another display, which has gaze focus. (c) The summary view for piling devices that displays a list of document titles, display orientations and application names when the pile of devices do not have body focus.

synchronized to a server. By aiming the tablet at any target monitor and tapping the "Connect to Display" button to confirm her intention (hand focus), the user will see user-annotated paragraphs displayed in a list view on the target monitor (Fig. 14b).

4.5 Piling

The last set of interaction techniques explored the opportunities that can be afforded when a system is able to determine that devices are not only physically located near each other, but also physically pointed towards each other. As the sorting and piling of documents and objects are common behaviors, bootstrapping such natural tendencies with digital devices will allow for several novel interactions - especially when combined with user-device information from hand, gaze, and body focus.

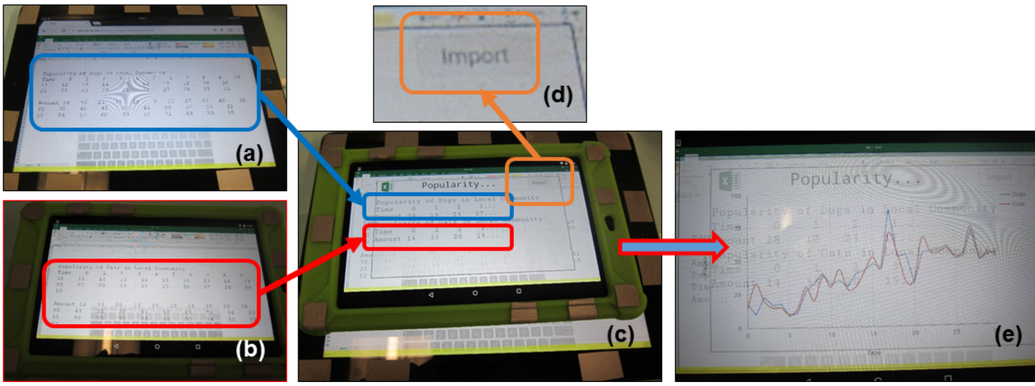

Fig. 16. Illustration of the *piling* state with the Piling Hub. (a, b) Two tablets containing different spreadsheets are piled atop one another. (c) When piled, the top tablet shows an integrated summary of the data from the pile. (d) With a tap of the button "Import", (e) the data is plotted using a unified visualization.

4.5.1 Piling Portal. As OmniDesk can detect whether devices are physically piled on top of each other, this information can be used to present interfaces that harness the thematic nature or links between devices. With the Piling Portal, the topmost device in a pile acts as a 'portal' to the other underlying devices. By gazing at the topmost device, the user can access information and data stored on all the piled devices without needing to physically separate and look at each individual device. Thus, the pile itself acts as an implicit interface to view the collection of information on each device. The topmost device, for example, could display all the applications that are running on the devices in the pile (Fig. 15a). By integrating hand focus with the Piling Portal, it is possible to enable navigation through the pile of stacked devices. The user can swipe through documents on the topmost tablet and make a selection, while the UI will highlight the application icon, theme color, and content (Fig. 15a). If the user looks at another monitor (gaze focus), the document navigation UI will be projected to the target device that has gaze focus (Fig. 15b). When the piled devices do not have body focus, the topmost device can display a list of documents using the Summary View technique, highlighting titles, display orientations, and application names (Fig. 15c).

4.5.2 Piling Hub. Another way that the piling of devices could be harnessed is by utilizing the topmost device as an aggregator, rather than a portal to the other devices in a pile. With the Piling Hub technique, whenever the topmost device receives gaze focus, it can reveal a summary of all the notifications or information contained on each device. Because hand focus indicates to the system that the user wishes to interact with all the contained information, it presents it in a manner than can be manipulated, rather than simply glanced at or navigated through.

If several piled tablets contain different spreadsheets, for example, the user can press the "Import" button displayed on the topmost device to generate an aggregated visualization of all the data in the pile (Fig. 16e). Such interaction adds physicality to the spreadsheet data and allows it to be combined and recombined by adding or removing devices from the pile via hand focus. Like the Piling Portal, gaze focus only presents information that can be viewed, whereas gaze and hand focus allow information to be manipulated and navigated.

We can further integrate the information attained from hand and gaze focus for other uses. After generating a new figure (Fig. 16e), we can move it into a document on another device by touching the figure (hand focus) and looking at the destination device (gaze focus; Fig. 17a). This action causes a pie menu to appear on the target monitor (Fig. 17b). The default option is set to "Cancel

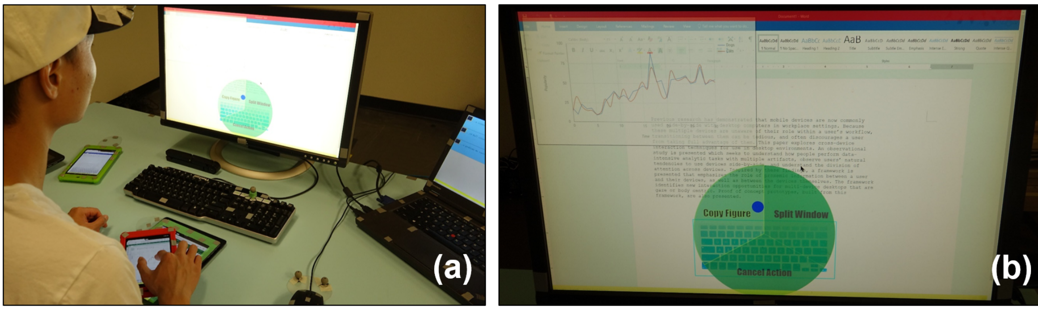

Fig. 17. Extending the Piling Hub with *gaze* and *hand* focus. (a) The user copies the figure by touching the tablet (hand focus) and (b) drops it on the target monitor (gaze focus) by selecting the "Copy Figure" option on the tablet (temporary "trackpad"; hand focus).

Action" in case the user accidentally touched the device or looked around while interacting with the device. Other options include moving the file to the target device and splitting the target display to compare the two files. The user can control the cursor on the target monitor by sliding her finger along the display of the source tablet, which has been temporarily converted to a trackpad.

4.6 Discussion

Guided by SMAC, the aforementioned examples demonstrated the unique opportunities that are afforded when one considers the role of user- and inter-device proxemics during the design of multi-device interaction system. The interaction technique exploration with the OmniDesk system demonstrated the power of considering both dimensions independently, as well as in combination with one or more other dimensions (Table 4). Techniques, such as the Gaze Copy and Adjustable Reader, explored the increased utility of attending to multiple user-device foci and examined the increased flexibility this affords users. In addition, techniques such as the Universal Keyboard and Digital Labels demonstrated how not only are the inter-device states themselves important, but so too are the transitions between these inter-device states. It is in such transitions that opportunities for enhanced and sustained feedback become possible.

Further, by looking at how prior work on cross-device interaction fits into SMAC, we can see that prior work focuses almost exclusively on the combination of (i) hand focus and (ii) either the broadcasting, expanding, or aiming state (Table 4). Interestingly, the use of body focus has largely gone unexplored. As our techniques have demonstrated, body focus is a useful indication of which devices that a user may want to use over a longer period of time because it is very unlikely that she may want to use devices that require her to hold an unnatural body posture. Even though the body has many more space restrictions placed on it in a desk-centric environment compared to mobile environments, there are many unconscious, rich cues that are still useful for developers to bootstrap.

Prior work has provided feasible solutions to enable wearable gaze tracking [50, 93] and demonstrated the benefit of gaze input for both pointing and context switching tasks [22, 107]. By combing gaze focus with other user- and inter-device dimensions in our model, more interactions become possible, as demonstrated in the OmniDesk system. Though improved sensing technology could provide more accurate results, unintentional input caused by accidental gaze switching or body movements should be carefully considered throughout the design process. A few techniques were built into OmniDesk to overcome some unintentional input scenarios, for example, the Following Window technique allows users to peek at the current characters being typed while consulting

Table 4. Summary of the new and existing techniques categorized by the dimensions of our model, SMAC. The interaction techniques illustrated by the OmniDesk system are in bold.

	Orientation: Away		**Orientation: Towards**	
	Distance: Separated (*Broadcasting*)	Distance: Adjacent (*Expanding*)	Distance: Separated (*Aiming*)	Distance: Adjacent (*Piling*)
Focus: Hand	**Device Radar; Digital Labels;** *Conductor* [33]; *Gradual Engagement* [63]; *Orienteer* [19]; *Pebbles* [72]; *Relate Gateways* [26]; *SyncTap* [84]; *Tracko* [45]; *United Slates* [11]; *Weave* [14]; *XD-Browser* [75].	**Adjustable Reader; Expanded Keyboard; Expanded Trackpad;** *Conductor* [33]; *HuddleLamp* [80]; *JuxtaPinch* [76]; *Panelrama* [105]; *Pass-Them-Around* [60]; *Stitching* [39]; *United Slates* [11].	**Universal Keyboard; SyncNotes;** *Code Space* [9]; *CTAT* [21]; *GroupTogether* [66]; *Touch Projector* [8]; *PhoneTouch* [88]; *Pipet* [67]; *Proxemic-Aware Controls* [57]; *Proxemic Interaction* [4]; *Slurp* [108]; *Throw and Tilt* [18].	**Piling Portal; Piling Hub;** *PaperTab* [97].
Focus: Gaze	**Gaze Copy; Gaze Search;** *LookPoint* [22]; *MAGIC Pointing* [107]; *Pupil* [50]; *ViewPointer* [93]; *Gluey* [92]; *Gaze Positioning* [98].	**Enhanced Adjustable Reader**	**Following Window**	**Piling Hub**
Focus: Body	**Gaze Search; Summary View;** *Chair Interaction* [79].	**Enhanced Adjustable Reader**	**SyncNotes**	**Piling Portal**

another document, however the input-capture relationship has to be changed explicitly by hand focus (e.g., via the Universal Keyboard). As yet another example, while displaying pie menus on the target device with gaze focus, the default cursor is set to "Cancel Action". This helps prevent triggering an option by accident. These techniques are simple, yet effective mechanisms through which some unintended input situations can be mitigated.

It is also interesting that few interaction techniques have explored how to harness the physical closeness and the directed orientation of devices, i.e., piling. Even though piling behaviors are very common in paper-based environments, the lack of techniques transferring such behaviors to digitally infused environments is surprising. Although our piling techniques explored how gaze and body focus can improve interaction when multiple devices are stacked, there are still many opportunities to continue exploring how such inter-device proxemic properties can be used. The use of hand and gaze focus within the present exploration was focused on all or one device receiving attention, however, there are many unique possibilities when sub-groups of stacked devices are considered. There thus appears to be further opportunities for developers and designers to reconsider the meaning and intentions behind the interactions that our whole body, not only our hands, has with devices.

The current implementation of OmniDesk used a Vicon motion tracking system, however, this was only due to equipment availability. With the proliferation of security systems in office settings and advancements in sensing techniques [55, 56], the increased variety of sensors within devices themselves (e.g., NFC and proximity sensors), and the increased adoption of intelligent tabletop surfaces that can sense the objects placed on them (e.g., Microsoft PixelSense), it will not be necessary to use Vicon-style systems in the future. Although the current implementation of OmniDesk was limited to desktop scenarios (so as to limit the capture volume of the Vicon system, among other reasons), OmniDesk could be extended to other environments as well, for example, mobile users in co-working spaces. Such extensions of the work would require that there is some mechanism by which users and their personal devices could be identified when they drop in to join the public space.

5 EVALUATION STUDY

As SMAC proposes a novel way the users can conceptualize the attention and focus they provide to devices, it is important to understand how users themselves react and understand the tenets of user-device (i.e., hand, gaze, and body focus) and inter-device (i.e., broadcasting, expanding, aiming, and piling state) proxemics. Just as window focus, keyboard focus, and z-order in a traditional WIMP have meanings that go beyond the mere facilitation of directing input, the present evaluation study sought to understand the degree to which SMAC gives rise to a consistent, understandable, and usable mental model of interleaved interaction in multi-device environments. Further, the study also sought to develop an understanding of whether the techniques that were designed using the dimensions of SMAC were understandable and could be embraced by users. Therefore, a summative study was conducted to understand participants' views of the dimensions governing SMAC as they related to everyday desk-based tasks and situations. The research questions guiding this exploration included:

- Q1: To what degree is each dimension an appropriate conceptualization of user intentions and interactions, i.e., can each dimension be easily understood by users?
- Q2: Can users articulate the differences between these dimensions, in terms of the effects each would have on a system, what they themselves would be enabled to do, and the feedback the user would receive, beyond the particular interaction methods that were developed?
- Q3: Can users understand these dimensions to the degree that they would be able to extrapolate from them to create new techniques or apply them to other scenarios?

To answer these questions, an interview-based study was conducted, wherein the techniques were presented to participants, organized by each dimension, and participants were asked questions to probe their understanding of the dimensions in our interaction model. By conducting a qualitative study, we sought to evaluate the value of SMAC as an understandable mental model rather than on the performance or efficiency of the system. As it is important for users to understand mental models of devices before they use them [52], the study was designed to be informal in nature to gauge how well participants could learn the model and its dimensions.

5.1 Participants

Eight participants (P1 - P8) who were not experts in the fields of Human-Computer Interaction or Information Technology (i.e., not working in the IT industry; not a student from Computer Science/Computer Engineering, etc.) but were regular users of laptop/desktop computers were recruited to participate in the study (2 male; *Mean* = 21.6, *Range* = 19-26 years). By excluding technology experts from our participant pool, we could investigate whether regular end-users, who don't commonly view interaction and computer usage as following specific models or classical concepts, could understand the dimensions and goals of SMAC, while also collecting valuable feedback on the design and the practicality of the techniques.

Table 5. Dimensions and techniques that were used within the evaluation study to probe participants' understanding of user-device and inter-device proxemics.

	Concept	Demonstrated Technique #1	Probed Technique #2
Introduction	Window Focus	Mouse	Keyboard (Alt-Tab)
User-Device Proxemics	Hand Focus	Mouse	Keyboard (Alt-Tab)
	Gaze Focus	Gaze Copy	Following Window
	Body Focus	Summary View	Gaze Search
Summary #1	User-Device Proxemics	Gaze Search	Following Window
Inter-Device Proxemics	Aiming	Remote Controller	Universal Keyboard
	Broadcasting	Device Radar	Digital Labels
	Expanding	Expanded Keyboard	Adjustable Reader
	Piling	Piling Hub	Piling Portal
Summary #2	User- and Inter-Device Proxemics	Enhanced Adjustable Reader	Piling Hub

One participant used multiple monitors for daily work, while two others had more limited experience using multiple monitors. None of the participants had experience using eye trackers or gaze-based systems. Participants were provided with a $20 CAD honorarium for participating in our one-hour interview.

5.2 Procedure

First, demographics and daily computer usage information were collected through a pre-study questionnaire. To prime participants for conversations about the concept of a *state*, we first introduced the WIMP concept of "Window Focus" by probing how participants would switch between two opened documents if they need to make a change in one of them. After discussing their answer, we further explained the concept of *focus*. An alternative solution was demonstrated using 'Alt-Tab' shortcut common on Windows OS to ensure that participants understood the concept of focus.

As extensions of this basic concept, we then probed how well participants could grasp the concept of user-device proxemics. We explained that the goal of our work was to extend the notion of window focus from a single device to an entire environment of tracked and connected devices. Participants were then shown various OmniDesk techniques and were invited to try out the gaze and body focus features. After this was complete, we explained the concept of hand focus and demonstrated it using previous examples of how to change window focus. We then asked probing questions to understand what the participant thought hand focus meant, and what the system did in response to hand focus. These simple questions about hand focus were used to introduce the way of considering the role of hand input in focus-switching to the participants.

A more complicated *demonstrate-and-probe* format was used for gaze and body focus, in which we presented OmniDesk techniques instead of showing the window focus example. A summary of the interaction techniques used in the study is listed in Table 5. For example, to study gaze focus, we first showed participants how the indicators on the UI would change when they switched their gaze focus between different devices. Then we demonstrated the first technique, i.e. Gaze Copy, to the participant, by showing them how this technique worked and answering any questions they had about it. Then we presented a different technique expressing the same dimension to the participant, i.e. Following Window, and asked specific questions about the second technique, including what the system did when the participant switched her gaze focus, and why the Following

Window technique worked the way we presented it. Participants were then asked subjective probing questions, including what they thought gaze focus meant, what other examples or scenarios of where gaze focus might be useful, and their general comments of presented examples. A similar demonstrate-and-probe format was used for all remaining techniques.

After presented with hand, gaze, and body focus, participants were then asked to explain what it meant to them when they were holding/touching a device (i.e., hand focus), looking at a device (i.e., gaze focus), and turning towards a device (i.e., body focus) and when they would do each of these activities, as a summary of the user-device proxemics section. A similar demonstrate-and-probe format was used in this session to present the techniques that used multiple foci. The researcher would summarize each dimension to the participants and then provided a one-minute break after this session.

Next, the study explored the concept of inter-device proxemics, using the same format as that used for probing gaze and body focus. This portion of the interview started with questions on the aiming state (as it is most akin to the metaphor of a TV remote controller), and then moved on to the broadcasting, expanding, and piling states. Finally, participants were presented with, and discussed, examples of inter-device proxemics and interaction techniques that were fusions of inter-device and user-device proxemics. They were also encouraged to ask questions and propose new ideas and comments during the interview. Although the methodology does not follow a traditional usability study, it encouraged exploration and reflection (via the examples), as well as knowledge extension (via the questions which tested how well they could apply the principles they had learned to other scenarios). Details of the questions asked during the interview are provided in Appendix A.

5.3 Data Collection

Each participant took between 45 to 60 minutes to complete the study. All study sessions were audio recorded and a total of 6.79 hours of audio was analyzed via open coding. The answers from participants were analyzed to determine their general understanding of a given dimension, the difficulties participants had in understanding certain aspects of the dimensions, and the practicality of the techniques and the overall interaction model for regular end-users. Participants' comments and suggestions on extensions of the dimensions to new scenarios were thematically organized by topic. Qualitative data analyses are provided for each dimension.

5.4 Findings

For readability, the results of the study first present the findings related to the user-device foci and techniques and are followed by those relating to the inter-device states and techniques.

5.4.1 Hand Focus. As hand focus is most akin to the traditional focus metaphors users employ on desktops, laptops, and smart phones, it is unsurprising that most participants (seven of the eight) caught onto this concept quickly. The examples of using mouse click and 'Alt-Tab' keys to switch window focus served as good illustrations to the user of the concept of 'focus'. Responses to *"What does it mean if this device has my hand focus? (The researcher is holding a tablet)"* ranged from *"This one is the only one that is active at the moment"* (P2) and *"You can use your hands to manipulate what you want to search, open ..."* (P6) to *"You are going to focus on that device now" (P4) and "You are going to switch applications and use it"* (P5). It thus appears that participants viewed the focus that the hand can provide to devices in terms of activation, attention, and input - tenets of interaction that indicate immediacy and intent (Q1).

The one participant who did not initially understand the concept of hand focus just considered it as simple 'singular focus' and was unable to articulate anything more about the attention or supported actions that were possible when one held or touched a device. We note that this participant did not have problems understanding any of the other concepts in the study.

Participants started to understand the roles of *hand focus* in the given examples, and this 'introducing' session also inspired the participants and encouraged them to consider the various foci in the following sessions.

5.4.2 Gaze Focus. All participants could apply the Gaze Copy technique to a different task after learning the first example (i.e. copy the application from the monitor to the tablet; see Appendix A), and all participants also used this knowledge to explain the Following Window technique (Q3). Most of the responses given by participants to what it means when you change what you are looking at revolved around specific tasks and applications, *"[You switch your gaze] if you want to copy something, or refer to something"* (P4), *"Looking at different materials" (P6), and "Check message, notifications, and song list"* (P7). Interestingly, these comments revealed that participants thought of gaze as part of information seeking and retrieval, and expected feedback to visualize the result of gaze-based actions. Thus cross-device systems should provide instant visual feedback to reassure the user that her gaze switch has been detected, and further information of the gazed device should be provided to support rapid information retrieval.

Most of the participants (seven of the eight) came up with new ideas on how gaze focus could be used, which was surprising given their limited experience with gaze-tracking systems. Four participants talked about improvements to existing applications, such as showing a floating window while Skyping (Q3), whereas three other participants suggested techniques that employed other peripherals, such as executing voice commands on a gaze-focused device or moving the mouse cursor with gaze focus (Q3). P8 extended this further by proposing the use of gaze focus to control fans and lights, but he also mentioned it would be a problem to accidentally turn the fans on and off while looking around. This participant's ability to recognize both a new extension of gaze focus but also the potential faults in such an approach reflects that he understood the principle of gaze focus, which treats the user's gaze as an indicator of attention and feedback, but also understood that another source of information (i.e., hand focus) would be needed to confirm input and prevent false activation (Q2).

These generated ideas and collected comments demonstrate that participants were able to understand and use gaze focus, echoing studies by Dickie et al. [22], but extend prior work by demonstrating how users can understand gaze focus to be useful within the context of desk-based environments and tasks.

5.4.3 Body Focus. Three participants were able to understand the roles of body focus in the Summary View example. While being probed with the second technique, Gaze Search, only one participant was able to determine that search results were ordered by relevance based on body focus (P7). P7 suggested that an extra check should be included while using body focus in case of users' random movement in the chair. This 'check' is already reflected in our model, as body focus is an indicator that affects the user interface but does not control it, and body focus needs to work in combination with other foci. Though seven other participants mentioned several factors that might affect the search results, such as relevance and publishing date, they didn't connect the concept of body focus with the relevance of the document to the current task. P3 commented that she did not think about differentiating documents on her desk by body focus because *"I will put everything in front of my body anyways"*. To this user, body focus was already implicitly applied and was an unconscious, rather than conscious, action to be taken.

After further explanation of body focus and answering their questions, six participants understood and proposed new ideas of using body focus. Two participants suggested using body focus to turn on/off the screen or sleep mode to save battery, and other four of six considered using body focus for special activities such as cooking or exercising (Q3). This result revealed that users are not familiar with using body focus as a factor that affects their workflow, as it is an implicit, unnoticed byproduct of their natural behavior (Q1).

5.4.4 Summary of User-Device Proxemics. When participants were encouraged to holistically consider these three foci, they talked about them in reference to the techniques they experienced. Six participants were able to describe them clearly and highlight their differences, such as *"Hand focus means you are concentrating on what you are holding; you look at that device if you want to remember something, or refer to something; body focus is the same [with gaze focus] but less focused, and hand focus is much more particular [compared with body focus]"* (P4), and *"Hand focus has most attention, gaze is less, and body is the least"* (P5). One participant also summarized a workflow of using multiple foci: *"[hand focus] You want to look at something in details, or searching for things you really want to work on; [gaze focus] You have already decided which one you are working on, and gaze focus pays attention to the task you are doing; [body focus] When you want to finish the current task and switch to look for other stuff and it will be a big change"* (P1).

These descriptions demonstrate that participants were able to understand the different dimensions, thus confirming Q2, and the descriptions that were given follow very closely to our own definitions of SMAC, which were used in this article but not explicitly described to participants using such language. Further, six participants explained the Following Window technique in a formal way that made use of the dimensions in our model (Fig. 1), demonstrating a further understanding of these concepts (Q1). Four participants were able to not only explain the dimensions, but also extend them to include other modalities, such as foot interaction or voice commands (Q3). Thus, for these participants, SMAC was a useful conceptual model that they used as a base upon which to think about other ways multi-device environments could be enhanced.

5.4.5 Aiming State. Unsurprisingly, the aiming state was easily understood by participants, due to the analogs it shares with traditional remote controllers. After learning about the Remote Controller, five participants could apply the aiming state to the keyboard, i.e., the Universal Keyboard, however two other participants commented that they didn't naturally think about moving the keyboard to aim at other devices, likely because they viewed the keyboard as a stationary, static device (Q1). P6 also expressed a usability concern, wherein *"You would have to have good sensitivity, because these devices are placed quite close together, they have to be very accurately picked up whether your gaze is on it or not"*, revealing the importance of real-time feedback for successful, or even available, connections (e.g., the gray and blue keyboard icons in Fig. 12) to facilitate better experiences.

Seven participants came up with new ideas for the aiming state, such as aiming tablets to control lighting or heating, and aiming to a hard drive to back up the data (Q3). Each of these ideas would require the user to set up connections and perform actions from the 'controller' side, which illustrates an understanding of the concept of the aiming state from SMAC. Their comments and ideas in this session suggested that the aiming state should be used to make changes to the system, as people do so with a TV remote controller.

5.4.6 Broadcasting State. In terms of the broadcasting, three participants commented that devices could not only share their location (as in the Device Radar), but also share content and functionality, as in the Digital Labels technique. Though the broadcasting state is the basic transition state in our model, i.e., all devices are initially separated and pointed away from each other, participants had trouble noticing its existence due to their abundance of experience using isolated devices. For example, P7 answered that *"They are syncing because the user is transferring information between devices"*, ignoring the inter-device relationships that existed between the device and believing that the user would control everything. In other words, the broadcasting state was providing information about surrounding devices to the back end of applications, without being noticed by the user.

We do note that four participants proposed new ideas using the broadcasting state, such as synchronizing changes in a document during group work, and quickly sharing files or links among local devices (Q3). These results thus suggest that broadcasting, like body focus, may be best viewed as an implicit state that enables the user to accomplish tasks, without explicitly requiring her to perform overt actions (like the other states require). It could thus help the user smooth the transition between tasks, activities, and other states.

5.4.7 Expanding State. Although many past research projects have explored expanding displays, our participants were not familiar with this functionality; however, they quickly grasped it. After observing the Expanded Keyboard technique, five participants could explain the main concepts behind the Adjustable Reader technique, i.e., devices in the expanding state act as one connected device because they are close to each other (Q1). Because this technique was inspired by people's reading behavior, where devices are close to each other and connected *"like a real book"* (P8), participants easily understood the importance of adjacency and expansion. One interesting observation was that P7 believed the expanded devices were synchronized to display continuous pages because *"You put them together and sync them"*, which, unlike the broadcasting state, but similar to the aiming state, was the result of the user intentionally performing an action, i.e., giving the device hand focus. In general, we can infer that the concept of inter-device states, which are manipulated by the distance and orientation between devices, rather than the participant's actions, were initially framed within participants' minds as the results of *their* actions not the *device's* actions. This is still a welcome finding, as it suggests that participants understood that all the devices were linked and that their actions were carried across devices, not isolated on each device.

However, even though they initially took a user-centric view of these states, they were able to generate new device-centric ideas using this state (Q3). Five participants suggested improvements for a reading software, such as adding a separate view of notes or references when comparing documents across expanded devices. Three participants wanted to apply the expanding state to other devices including cameras, gaming devices, and external monitors. Each of these ideas demonstrate the usefulness of the expanding state in a variety of environments, and participants' ability to understand the concept of the expanding state, i.e., devices are tightly connected to continually exchange data within the group until they are no longer connected.

5.4.8 Piling State. During the exploration of the piling state, six participants were able to quickly explain the concept of piling when shown different content (e.g., two Excel datasheets; Q1). They described the role of the devices in reference to the z-order of the devices, e.g., *"When it is on top of the other, it shows something underneath"* (P2), *"Indicating there are other devices underneath"* (P3), and *"[Because] they have physical contact, whatever is on top you are navigating knows that if should show you preview of what's on both devices"* (P6). These comments demonstrated that, unprompted, participants immediately realized that the topmost device was an aggregator to the resources underneath and should display summary information with interactive views, thus forming a portal for the entire pile.

With this mental model, two participants immediately saw cost and organizational limitations with piling, i.e., *"I don't think people will use multiple tablets because they are expensive"* (P7) and *"I would rather lay them around, so I don't have to stack them [unless] I need to save some space"* (P6). Although tablets may become cheap and ubiquitous enough to overcome the cost concern, as suggested by P6, designers and developers may need to work to reframe the notion of 'device' such that it does not imply a singular, dedicated computer that performs its own isolated task. Instead, a device would be part of an interconnected network, and temporarily display content and accept input such that networks of them could be piled, shuffled, and reorganized as needed.

All participants proposed new ideas using the piling state, such as sorting documents, comparing documents or images, and merging notes from different co-workers (Q3). Among the ideas they mentioned, most of them were thematically related to addressing existing challenges with handling physical documents, suggesting that participant's mental models viewed the tablets as documents, rather than digital devices. Supported by the light-weight paper tablets in the future (e.g., PaperTab [97]), this finding echoes to Weiser's vision of spreading 'Pads' around the desk just as people "spread out papers" [102].

5.4.9 Summary of User- and Inter-Device Proxemics. During the summary session of the interviews, most participants (seven of eight) were able to identify the dimensions of SMAC that were utilized within one of the example techniques we demonstrated, and five of them even identified the model dimensions in both techniques (Q2). Since the two examples in this session combined multiple dimensions from the user- and inter-device proxemics, these results revealed that participants were able to quickly learn the concepts from the previous sessions and apply the interaction model to more complicated interactions.

Most participants (seven of eight) were able to come up with new ideas using the concepts in SMAC, including five of them who proposed extensions to existing techniques (Q3). Two participants even combined the user- and inter-device relationships together, using gaze focus and the broadcasting state to move files from a phone to a TV, and using gaze focus and the expanding state to copy figures from a PDF file to their notes. P7 suggested an extension to the piling state, whereby a sub-state could detect if devices were being flipped over and placed together, as opposed to being piled on top of each other, all facing up.

All of these suggestions are inspiring, as they suggest that, even in the short duration of time that participants were exposed to this model, participants were not only able to understand each dimension and identify which one was used in a given technique, but also provide other examples of their use and, in some cases, extensions that employed other modalities or more fine-grained information from a multi-device desk-centric environment than what we had suggested.

5.5 Discussion

In general, this study demonstrated that participants were not only able to understand the concept of window focus, but also understand and extrapolate this concept to multi-device environments. Although these participants are not normally attuned to thinking about the models or metaphors that govern the actions they perform with their devices, their comments indicated that, they do unconsciously make use of an internal mental model to govern their interaction. Through the demonstrate-and-probe methodology that was used, we found that participants were able to build on, and extrapolate from, this model to integrate the new SMAC dimensions that we proposed into their mental model. The ability of participants to then use these dimensions, which they had never seen before and only had a short duration of time to become familiar with, speaks to the appropriateness and power of using such an interaction model to describe, design, and build new techniques for use in multi-device desk-centric environments.

The results from the user-device proxemics portion of the study demonstrated that participants were quick to understand foci that demanded their explicit interaction, i.e., gaze and hand, but were slower to understand those that related to cues that were subliminal, covert, and unintended, i.e., body. As our brief discussion of body focus indicated, because participants had not thought about their body as an indicator of their intentions, they initially had a difficult time determining what turning one's body towards a device would cause a system to do or enable them to do. As looking at and touching objects within our world are natural ways to give attention to objects, we often use these movements to indicate temporary fixtures of attention. As a result, it is not surprising that participants were easily able to grasp these concepts, generate novel techniques that make use of them, and think of other modalities that could be repurposed in a similar manner (e.g., voice, foot tapping). With body focus, if a system is aware that the orientation of the body indicates a general, more long-term area of interest, it may simply be enough for the system to initially use body focus as a coarse information indicator within a focus-based multi-device system, and later refine the focus state with more fine-grained information from gaze focus and hand focus. Although users may not have attended to, or understood why the Gaze Search technique returned the ‘right’ document at the top of the list, a crucial question to ask is, do they need to? If their body indicates general zones where attention and focus will be directed for long periods of time, perhaps body focus might also be best leveraged with techniques that utilize it as an implicit cue to improve usability, without considering or teaching the user that it can be an explicitly controlled input modality.

In looking at the inter-device proxemic state results, it was interesting that participants viewed the inter-device proxemic states in reference to themselves and their actions, rather than the current location and orientation of the devices. While this may be a by-product of the methodology, which introduced the concept of focus to the participant and had them use this notion to understand that the user was causing changes within an environment, we do not believe that this is the entire explanation. As the broadcasting results demonstrated (Section 5.4.6), users internalize the outcomes and feedback from a system as reactions to their actions. In some cases, such as with the aiming state, they also appear to temporarily extend their ‘power’ to devices themselves. As the aiming state is one of the only states where the user holds a device to perform an action (e.g., a mouse or remote controller), the sheer act of physically holding a device versus moving it along a table or surface may help the user differentiate between devices that the user has complete control over and those that they do not. This is further evidenced by participants using terms during the study that referenced the devices themselves when describing the piling and expanding states, versus their own actions when describing the aiming state.

6 DESIGN IMPLICATIONS

Based on our exploration of user-device and inter-device proxemics, many implications and future challenges became apparent. Though we focused on the cross-device interaction in a desk-based environment, SMAC could be extended to many other fields, including smart homes, multi-user collaboration environments, and mobile computing. For example, the design space of inter-device proxemics lends itself to commercial smart furniture and could help bridge the barriers amongst prevailing smart home devices. In addition, existing user- and inter-device proxemics could be accompanied with inter-user proxemics [66], to support multi-user collaborative tasks. Wearable devices also provide for finer-grained user-device proxemics, tracking multiple parts of the body such as the finger, head, arm, and leg. On the other hand, there are still challenges to be addressed. The solution we used within the OmniDesk prototype system is far from perfect, as it relies on motion tracking system and tagged devices. It is also challenging to obtain the precise location of individual artifacts in the workspace, especially when they are placed in piles or occluded by each other.

Here we present some findings from our study and prototype implementation to illustrate how our interaction model can contribute to the community and those exploring multi-device ecologies.

6.1 Importance of Materiality

Although the paperless office was foreseen in previous research, information workers still rely, and prefer, paper documents. Paper can be moved around and are not restricted by the boundaries of a device, even when we have one as large as a tabletop or a wall display. By encouraging users to handle and manipulate their documents, users have more flexibility and a different experience compared to dragging and organizing virtual files on the display [43]. However, paper is less interactive compared with other digital devices. When the user prints a paper document from a digital version, the connection between the two artefacts ends immediately. The printed document is unaware of any updates on its digital version, and the computer is uninformed for annotations or notes on the physical documents. That is the reason why devices should be connected and broadcast their resources, locations, and postures to others. In this way, different resources are treated as independent objects in a network, instead of a group of virtual icons on a large monitor. As a result, the relationships between the devices indicate the user's intent, and provide opportunities for richer user- and inter-device interactions.

To investigate a proper implementation of this vision, we focus on the desk-centric environment with multiple tablets in our prototype. Learning from the materiality of paper, tablets could predominate by deriving natural operations such as piling on another, spreading on the desk, holding with the hands, or augmenting them with inter-device proxemics and users' foci. With the desktop computer still being the primary device, tablets should accompany the desktop to replace the current printed documents, when appropriate, and provide flexible interaction methods with the touch display and built-in sensors.

6.2 Unintentional Input from Body Gestures

With more sensors and assistive algorithms integrated into working environments, it is essential that systems are robust to unintentional input. When systems are not able to identify intentional versus unintentional input, failures could include the user accidentally touching the surface while holding a device with hands, or the user quickly looking at a new stimulus while wearing an eye tracker. On the one hand, tracking the user's body gestures enables more natural interactions to assist with traditional keyboard and mouse usage, which can be used to help predict the user's target display, the user's preferred input device, and the resources that the user may need. However, tracking could also discourage the user, which could lead him sit on a chair adroitly so as to not trigger any 'strange' events. As a result, 'coarse' attention indicators (e.g., gaze and body focus) should be confirmed by hand focus before actual changes are made within the system. For example, in the Following Window, rather than using gaze to switch output devices while typing, the characters being typed are forwarded as a hint. This ensures that the original input capture relationship can be maintained until the user explicitly switches output devices with her hand.

In addition, it should be easy to recover from erroneous state changes, so that the user will not feel nervous while moving around naturally. In the OmniDesk system, we set the default selection target to the "Cancel Action" button on the gaze-directed menu in order not to distract the user. The gestures in a practical system should rely on observations from the user's current workflow or involve user-defined gestures to improve the learnability of the interaction techniques.

6.3 Multiple Devices of the Same Type

Vision-based tracking techniques are useful to track the 3D location of multiple devices. However, they lack precision when multiple devices of the same type are piled in a stack or grouped close to each other. CapCam [104] tracks the physical locations of multiple phones on a tabletop by analyzing the encoded signal from the phone's rear camera. SurfaceLink [27] supports gestures between the devices on a shared table using acoustic sensing. HuddleLamp [80] combined RGB images with depth data to improve the recognition of hand gestures and devices. Though these systems focus on device tracking on a shared 2D surface, they could be used to help to improve the robustness of OmniDesk's 3D tracking system and contribute to an ideal digital workspace.

If devices are pervasive and available anywhere in the office, they could also be identified and assigned different roles [102]. Given a large collection of tablets, which may be commonplace in future work-based environments, some may act as storage media, while others may become tools like whiteboards or remote controllers. For this to occur, however, users must be trained to change their view of tablets as 'computers' to 'documents'. As evidenced in work by Plank et al. [78], people's "legacy bias" results in participants constantly changed the function and content of the 'computers', instead of assigning the tablets with a long-lasting role as a 'document'. To address this challenge, designers should provide users with the feelings of holding, reading, and writing on a piece of 'paper', thanks to the light weight tablets (e.g., PaperTab [97] or Sony Digital Paper [94]). Designers could also reduce the mental demand of using multiple tablets by simplifying the functionality supported within current tasks and improving the learnability of mobile applications [59].

6.4 Multiple Users

By extending the design considerations in our single-user environment, the SMAC model could be used to define useful behaviors when multiple users focus on a single device, or multiple devices belong to different users trying to communicate. For example, if Bob is walking towards Alice's desk with his tablet and wants to show her a figure, while he is doing this, the devices on Bob's desk should lock after they have lost his attention (body, gaze, and hand focus). When he arrives at Alice's desk, both his tablet and Alice's desktop monitor should adjust their content when they detect that a new user is gazing upon them (gaze focus). With Alice's approval, her monitor should also broadcast its content to a larger, nearby touchscreen device, so that Bob can interact with it (hand focus). This example is similar to the interconnected tablets work by can Hinckley et al. [38] and Chen et al. [11], which advocated for supporting collaboration in office environments, but also demonstrates how the tenets of the SMAC model can be useful to define the boundaries of collaboration, interaction, and privacy. Future work should consider how SMAC can be used to broaden the scope of current workflows and collaboration from single personal workspaces to an entire collection or office of workspaces.

6.5 Intimate Devices

As wearable devices become prevalent, they can provide more input channels to a system in an intimate way. Some devices intend to track the movement of the user's arms, hands, and fingers accurately. Smartwatches support foreground gesture interactions on the surface, and also provide background sensors naturally tracking the user's hand [12, 41]. Smart rings could be embedded with motion sensors and vibration motors to support input and output at a more private level.

Smart glasses have also become popular with commercial products like Google Glass and Microsoft HoloLens, enabling a mobile private 'display' to follow the user's head. Prior work also proposed mobile solutions for eye-tracking, such as ViewPointer [93] and Pupil [50]. One drawback of the OmniDesk prototype was that it tracked the user's head and used that as a proxy for gaze instead of tracking the user's eyes. For an initial exploration, this abstraction was usable, but in practice it is not as precise as using an eye tracking because the user could switch her gaze focus without moving her head. However, it could be a benefit to introduce the concept of head focus (for some systems) and explore the utility of gaze focus and head focus together, because each convey different user intentions. For example, if a user faces her colleagues while looking down at her tablet, this could imply that she may want to search for something on her device to share with others. As another example, a user may face her primary monitor (head focus), while at the same time glancing at the secondary monitor next to it. Since it is not comfortable for a user to maintain this posture, such behavior could be used to indicate a temporary switch in attention instead of a permanent one. The system could then respond based on the current content and application history. Though we only mentioned using the focus of hand, gaze, and body in the user-device proxemics, these intimate devices and sensors provide increasing opportunities to understand the user's intent by detecting the focus of finer-grained movement and behavior, such as that from the finger, head, arm, or leg.

There is also a trade-off between wearing multiple devices on different parts of the body and installing a motion capture system in the environment. Wearable devices are flexible and could be customized because they are considered personal and private, but systems relying on them would fail to respond to other users in the room who do not have on-body sensors. There may also be fatigue issues if users need to wear multiple wearable devices to achieve the same level of tracking as that currently possible with motion capture system. On the other hand, motion capture systems can track multiple users in the environment, but are limited to certain physical spaces and are not able to respond to user activities outside the space. With more efforts to integrate intimate devices into the surrounding smart environment, we are moving closer and closer Weiser's vision of invisible computing without barriers between devices [102].

7 CONCLUSION

This work has explored a new way to conceptualize the usage of multiple devices in a desk-centric environment. By integrating inter-device proxemics, i.e., device orientation and location, with user-centric dimensions including the user's hand focus, eye gaze, and body orientation, SMAC describes a novel way to design for multi-device environments. By focusing on the interplay between the cues a user *and* her devices convey, designers and developers can create rich, responsive systems that reflect both intentional input on the part of the user in terms of her interaction with devices (i.e., hand focus, gaze focus, aiming with devices), but also unconscious intentions that are a byproduct of her natural behavior (i.e., body focus, piling devices, etc.).

Using a proof-of-concept system, OmniDesk, we explored the design space of dimensions and opportunities SMAC affords. An evaluation study was conducted to demonstrate the comprehensiveness of our interaction model using regular end-users of desktop computers. The study findings uncovered that not only could users understand the dimensions of SMAC and identify which were utilized within an example interaction technique, but users were also able to extrapolate from the provided examples and suggested other modalities or sub-states that could be harnessed to improve multi-device desk-centric environments. The results of the study, in combination with SMAC itself, should serve to stimulate and guide future discussions on how to design interaction to ensure users with multiple devices in desk-centric systems have rich and seamless experiences.

ACKNOWLEDGMENTS

We thank Mingming Fan and Peter Hamilton for providing helpful suggestions, and the members of DGP lab for their valuable discussions.

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)*. 55–64. https://doi.org/10.1145/1753326.1753336

[2] Apple Computer Inc. 2005. Main, Key, and Inactive Windows. In *Apple Human Interface Guidelines*. Chapter 13, 344. http://www.multimedialab.be/doc/tech/doc_osx_hi_guidelines.pdf

[3] Sriram Karthik Badam and Niklas Elmqvist. 2014. PolyChrome: A Cross-Device Framework for Collaborative Web Visualization. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '14)*. 109–118. https://doi.org/10.1145/2669485.2669518

[4] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '10)*. 121–130. https://doi.org/10.1145/1936652.1936676

[5] Patrick Baudisch, Nathaniel Good, and Paul Stewart. 2001. Focus Plus Context Screens: Combining Display Technology with Visualization Techniques. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '01)*. 31–40. https://doi.org/10.1145/502348.502354

[6] Xiaojun Bi, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2011. Magic Desk: Bringing Multi-Touch Surfaces into Desktop Work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 2511–2520. https://doi.org/10.1145/1978942.1979309

[7] Olha Bondarenko and Ruud. Janssen. 2005. Documents at Hand: Learning from Paper to Improve Digital Technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. 121–130. https://doi.org/10.1.1.137.7259

[8] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. 2010. Touch Projector: Mobile Interaction through Video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. 2287–2296. https://doi.org/10.1145/1753326.1753671

[9] Andrew Bragdon, Rob DeLine, Ken Hinckley, and Meredith Ringel Morris. 2011. Code Space: Touch+ Air Gesture Hybrid Interactions for Supporting Developer Meetings. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '11)*. 212–221. https://doi.org/10.1145/2076354.2076393 arXiv:arXiv:1011.1669v3

[10] Tsung-Hsiang Chang and Yang Li. 2011. Deep Shot: A Framework for Migrating Tasks Across Devices Using Mobile Phone Cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 2163–2172. https://doi.org/10.1145/1978942.1979257

[11] Nicholas Chen, Francois Guimbretiere, and Abigail Sellen. 2012. Designing a Multi-Slate Reading Environment to Support Active Reading Activities. *ACM Transactions on Computer-Human Interaction* 19, 3 (2012), 18. https://doi.org/10.1145/2362364.2362366

[12] Xiang 'Anthony' Chen, Tovi Grossman, Daniel J. Wigdor, and George Fitzmaurice. 2014. Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 159–168. https://doi.org/10.1145/2556288.2556955

[13] Victor Cheung and Stacey D. Scott. 2015. Studying Attraction Power in Proxemics-Based Visual Concepts for Large Public Interactive Displays. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '15)*. 93–102. https://doi.org/10.1145/2817721.2817749

[14] Pei-Yu (Peggy) Chi and Yang Li. 2015. Weave: Scripting Cross-Device Wearable Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. 3923–3932. https://doi.org/10.1145/2702123.2702451

[15] Ming Ki Chong and Hans Gellersen. 2011. How Users Associate Wireless Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 1909–1918. https://doi.org/10.1145/1978942.1979219

[16] Ming Ki Chong, Rene Mayrhofer, and Hans Gellersen. 2014. A survey of user interaction for spontaneous device association. *ACM Computing Surveys (CSUR)* 47, 1 (2014), 8. https://doi.org/10.1145/2597768

[17] Irene Cole. 1982. Human Aspects of Office Filing: Implications for the Electronic Office. In *Proceedings of the Human Factors and Ergonomics Society Annual Meetings*, Vol. 26. 59–63.

[18] Raimund Dachselt, User Interface, Engineering Group, Robert Buchholz, and Security Group. 2009. Natural Throw and Tilt Interaction between Mobile Phones and Distant Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '09)*. 3253–3258. https://doi.org/10.1145/1520340.1520467

[19] David Dearman, Richard Guy, and Khai Truong. 2012. Determining the Orientation of Proximate Mobile Devices using their Back Facing Camera. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. 2231–2234. https://doi.org/10.1145/2207676.2208377

[20] David Dearman and Jeffrey S Pierce. 2008. "It's on my other Computer!": Computing with Multiple Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. 767–776. https://doi.org/10.1145/1357054.1357177

[21] Linda Di Geronimo, Maria Husmann, Abhimanyu Patel, Can Tuerk, and Moira C Norrie. 2016. CTAT: Tilt-and-Tap Across Devices. In *International Conference on Web Engineering*. Springer, 96–113. https://doi.org/10.1007/3-540-45144-7

[22] Connor Dickie, Jamie Hart, Roel Vertegaal, and Alex Eiser. 2006. LookPoint: An Evaluation of Eye Input for Hands-Free Switching of Input Devices between Multiple Computers. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '06)*. 119–126. https://doi.org/10.1145/1228175.1228198

[23] George W. Fitzmaurice, Azam Khan, William Buxton, Gordon Kurtenback, and Ravin Balakrishnan. 2003. Sentient Data Access via a Diverse Society of Devices. *ACM Queue* 1, 8 (2003), 53–62. https://doi.org/10.1145/966712.966721
[24] Luca Frosini and Fabio Paternò. 2014. User Interface Distribution in Multi-Device and Multi-User Environments with Dynamically Migrating Engines. In *Proceedings of the ACM SIGCHI symposium on Engineering interactive computing systems (EICS '14)*. 55–64. https://doi.org/10.1145/2607023.2607032
[25] Jörg Geißler. 1998. Shuffle, Throw or Take it! Working Efficiently With an Interactive Wall. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98)*. 265–266. https://doi.org/10.1145/286498.286745
[26] Hans Gellersen, Carl Fischer, Dominique Guinard, Roswitha Gostner, Gerd Kortuem, Christian Kray, Enrico Rukzio, and Sara Streng. 2009. Supporting device discovery and spontaneous interaction with spatial references. *Personal and Ubiquitous Computing* 13, 4 (2009), 255–264. https://doi.org/10.1007/s00779-008-0206-3
[27] Mayank Goel, Brendan Lee, Md. Tanvir Islam Aumi, Shwetak Patel, Gaetano Borriello, Stacie Hibino, and Bo Begole. 2014. SurfaceLink: Using Inertial and Acoustic Sensing to Enable Multi-Device Interaction on a Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 1387–1396. https://doi.org/10.1145/2556288.2557120
[28] Francisco J. González-Castaño, Javier Vales-Alonso, Miron Livny, Enrique Costa-Montenegro, and Luis Anido-Rifón. 2002. Condor grid computing from mobile handheld devices. *ACM SIGMOBILE Mobile Computing and Communications Review* 6, 2 (2002), 18–27. https://doi.org/10.1145/881978.882005
[29] S. Greenberg, N. Marquardt, T. Ballendat, R. Diaz-Marino, and M. Wang. 2011. Proxemic Interactions: The New Ubicomp? *Interactions* 18, 1 (2011), 42–50. https://doi.org/10.1145/1897239.1897250
[30] Jonathan Grudin. 2001. Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. 458–465. https://doi.org/10.1145/365024.365312
[31] Edward T. Hall. 1963. A System for the Notation of Proxemic Behavior. *American Anthropologist* 65, 5 (1963), 1003–1026.
[32] Edward T. Hall. 1966. *The Hidden Dimension.* Doubleday & Co, New York.
[33] 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)*. 2773–2782. https://doi.org/10.1145/2556288.2557170
[34] Robert Hardy and Enrico Rukzio. 2008. Touch & Interact: Touch- based Interaction of Mobile Phones with Displays. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI '08)*. 245–254. https://doi.org/10.1145/1409240.1409267
[35] Chris Harrison and Anind K. Dey. 2008. Lean and Zoom: Proximity-Aware User Interface and Content Magnification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. 507–510. https://doi.org/10.1145/1357054.1357135
[36] Doris Hausen, Sebastian Boring, and Saul Greenberg. 2013. The Unadorned Desk: Exploiting the Physical Space around a Display as an Input Canvas. *IFIP Conference on Human-Computer Interaction. Springer Berlin Heidelberg, 2013* 8117 (2013), 140–158. https://doi.org/10.1007/978-3-642-40483-2_10
[37] Ken Hinckley. 2003. Synchronous gestures for multiple persons and computers. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '03)*, Vol. 5. 149–158. https://doi.org/10.1145/964696.964713
[38] Ken Hinckley, Morgan Dixon, Raman Sarin, Francois Guimbretiere, and Ravin Balakrishnan. 2009. Codex: A Dual Screen Tablet Computer. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. 1933–1942. https://doi.org/10.1145/1518701.1518996
[39] Ken Hinckley, Gonzalo Ramos, Francois Guimbretiere, Patrick Baudisch, and Marc Smith. 2004. Stitching: Pen Gestures that Span Multiple Displays. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '04)*. 23–31. https://doi.org/10.1145/989863.989866
[40] Matthew Hong, Anne Marie Piper, Nadir Weibel, Simon Olberding, and James Hollan. 2012. Microanalysis of Active Reading Behavior to Inform Design of Interactive Desktop Workspaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. 215–224. https://doi.org/10.1145/2396636.2396670
[41] Steven Houben and Nicolai Marquardt. 2015. WatchConnect: A Tookit for Prototyping Smartwatch-Centric Cross-Device Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. 1247–1256. https://doi.org/10.1145/2702123.2702215
[42] Steven Houben, Paolo Tell, and Jakob Bardram. 2014. ActivitySpace: Managing Device Ecologies in an Activity-Centric Configuration Space. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '14)*. 119–128. https://doi.org/10.1145/2669485.2669493
[43] Samuel Huron, Yvonne Jansen, and Sheelagh Carpendale. 2014. Constructing Visual Representations: Investigating the Use of Tangible Tokens. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2102–2111. https://doi.org/10.1109/TVCG.2014.2346292

[44] Pradthana Jarusriboonchai, Thomas Olsson, Vikas Prabhu, and Kaisa Vaananen-Vainio-Mattila. 2015. CueSense: a Wearable Proximity-Aware Display Enhancing Encounters. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '15)*. 2127–2132. https://doi.org/10.1145/2702613.2732833

[45] Haojian Jin, Christian Holz, and Kasper Hornbæk. 2015. Tracko: Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '15)*. 147–156. https://doi.org/10.1145/2807442.2807475

[46] Wendy G Ju, Brian A Lee, and Scott R Klemmer. 2007. Range - Exploring Proxemics in Collaborative Whiteboard Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '07)*. 2483–2488. https://doi.org/10.1145/1240866.1241028

[47] Shaun K. Kane, Daniel Avrahami, Jacob O. Wobbrock, Beverly Harrison, Adam D. Rea, Matthai Philipose, and Anthony LaMarca. 2009. Bonfire: A Nomadic System for Hybrid Laptop-Tabletop Interaction. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '09)*. 129–138. http://dl.acm.org/citation.cfm?id=1622202

[48] Shaun K Kane, Amy K Karlson, Brian R Meyers, Paul Johns, Andy Jacobs, and Greg Smith. 2009. Exploring Cross-Device Web Use on PCs and Mobile Devices. *INTERACT 2009* (2009), 1–14.

[49] Amy K. Karlson, Brian R. Meyers, Andy Jacobs, Paul Johns, and Shaun K. Kane. 2009. Working overtime: Patterns of smartphone and PC usage in the day of an information worker. In *Proceedings of the International Conference on Pervasive Computing (Pervasive '09)*. 398–405. https://doi.org/10.1007/978-3-642-01516-8_27

[50] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. 1151–1160. https://doi.org/10.1145/2638728.2641695 arXiv:1405.0006

[51] Alison Kidd. 1994. The Marks are on the Knowledge Worker. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)* (1994), 186–191. https://doi.org/10.1145/191666.191740

[52] David E Kieras and Susan Bovair. 1984. The Role of a Mental Model in Learning to Operate a Device. *Cognitive Science* 8, 3 (1984), 255–273. https://doi.org/10.1207/s15516709cog0803_3

[53] David Kirsh. 1995. The Intelligent Use of Space. *Artificial Intelligence* 73, 1 (1995), 31–68.

[54] Ulrike Kister, Patrick Reipschl, Fabrice Matulic, and Raimund Dachselt. 2015. BodyLenses - Embodied Magic Lenses and Personal Territories for Wall Displays. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '15)*. 117–126. https://doi.org/10.1145/2817721.2817726

[55] Gierad Laput, Xiang 'Anthony' Chen, and Chris Harrison. 2016. SweepSense: Ad Hoc Configuration Sensing Using Reflected Swept-Frequency Ultrasonics. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI '16)*. 332–335. https://doi.org/10.1145/2856767.2856812

[56] Gierad Laput, Chouchang Yang, Robert Xiao, and Alanson Sample. 2015. EM-Sense : Touch Recognition of Uninstrumented , Electrical and Electromechanical Objects. In *ACM Symposium on User Interface Software and Technology (UIST '15)*. 157–166. https://doi.org/10.1145/2807442.2807481

[57] David Ledo, Saul Greenberg, Nicolai Marquardt, and Sebastian Boring. 2015. Proxemic-Aware Controls: Designing Remote Controls for Ubiquitous Computing Ecologies. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI '15)*. 187–198. https://doi.org/10.1145/2785830.2785871

[58] Anke Lehmann, Heidrun Schumann, Oliver Staadt, and Christian Tominski. 2011. Physical Navigation to Support Graph Exploration on a Large High-Resolution Display. In *International Symposium on Visual Computing (ISVC '11)*. 496–507. https://doi.org/10.1007/978-3-642-24028-7_46

[59] Rock Leung, Leah Findlater, Joanna McGrenere, Peter Graf, and Justine Yang. 2010. Multi-Layered Interfaces to Improve Older Adults' Initial Learnability of Mobile Applications. *ACM Transactions on Accessible Computing* 3, 1 (2010), 1–30. https://doi.org/10.1145/1838562.1838563

[60] Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2011. Pass-Them-Around: Collaborative Use of Mobile Phones for Photo Sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. 1787–1796. https://doi.org/10.1145/1978942.1979201

[61] Thomas W. Malone. 1983. How Do People Organize Their Desks?: Implications for the Design of Office Information Systems. *ACM Transactions on Information Systems* 1, 1 (1983), 99–112. https://doi.org/10.1145/357423.357430

[62] Marco Manca and Fabio Paternò. 2016. Customizable Dynamic User Interface distribution. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '16)*. 27–37. https://doi.org/10.1145/2933242.2933259

[63] Nicolai Marquardt, Till Ballendat, Sebastian Boring, Saul Greenberg, and Ken Hinckley. 2012. Gradual Engagement: Facilitating Information Exchange between Digital Devices as a Function of Proximity. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '12)*. 31–40. https://doi.org/10.1145/2396636.2396642

[64] Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. 2011. The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '11)*. 315–325. https://doi.org/10.1145/1979742.1979691

[65] Nicolai Marquardt and Saul Greenberg. 2015. *Proxemic Interactions: From Theory to Practice.* Morgan & Claypool Publishers. https://doi.org/10.2200/S00619ED1V01Y201502HCI025

[66] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-Device Interaction via Micro-mobility and F-formations. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '12)*. 13–22. https://doi.org/10.1145/2380116.2380121

[67] Bernt Meerbeek, Peter Bingley, Wil Rijnen, and Elise van den Hoven. 2010. Pipet: a design concept supporting photo sharing. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)*. 335–342. https://doi.org/10.1145/1868914.1868954

[68] Jérémie Melchior, Donatien Grolaux, Jean Vanderdonckt, and Peter Van Roy. 2009. A Toolkit for Peer-to-Peer Distributed User Interfaces: Concepts, Implementation, and Applications. In *Proceedings of the ACM SIGCHI symposium on Engineering interactive computing systems (EICS '09)*. 69–78. https://doi.org/10.1145/1570433.1570449

[69] Jérémie Melchior, Jean Vanderdonckt, and Peter Van Roy. 2011. A Model-Based Approach for Distributed User Interfaces. In *Proceedings of the ACM SIGCHI symposium on Engineering interactive computing systems (EICS '11)*. 11–20. https://doi.org/10.1145/1996461.1996488

[70] Microsoft Docs. 2018. About Mouse Input. https://docs.microsoft.com/en-us/windows/desktop/inputdev/about-mouse-input

[71] Diane Mizrachi. 2013. Individuality and diversity among undergraduates' academic information behaviors: An exploratory study. *International Journal of Knowledge Content Development & Technology* 3, 2 (2013), 29–42. https://doi.org/10.5865/IJKCT.2013.3.2.029

[72] Brad A Myers. 2001. Using Handhelds and PCs Together. *Commun. ACM* 44, 11 (2001), 34–41.

[73] Brad A Myers, Choon Hong Peck, Jeffrey Nichols, Dave Kong, and Robert Miller. 2001. Interacting At a Distance Using Semantic Snarfing. In *International Conference on Ubiquitous Computing*. Springer Berlin Heidelberg, 305–314. https://doi.org/10.1145/503376.503383

[74] Miguel A. Nacenta, Dzmitry Aliakseyeu, Sriram Subramanian, and Carl Gutwin. 2005. A Comparison of Techniques for Multi-Display Reaching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. 371–380. https://doi.org/10.1145/1054972.1055024

[75] Michael Nebeling and Anind K Dey. 2016. XDBrowser : User-Defined Cross-Device Web Page Designs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*. 5494–5505. https://doi.org/10.1145/2858036.2858048

[76] Heidi Selmer Nielsen, Marius Pallisgaard Olsen, Mikael B. Skov, and Jesper Kjeldskov. 2014. JuxtaPinch: Exploring Multi-Device Interaction in Collocated Photo Sharing. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices & Services (MobileHCI '14)*. 183–192. https://doi.org/10.1145/2628363.2628369

[77] Antti Oulasvirta and Lauri Sumari. 2007. Mobile kits and laptop trays: managing multiple devices in mobile information work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. 1127–1136. https://doi.org/10.1145/1240624.1240795

[78] Thomas Plank, Hans-christian Jetter, Roman Rädle, Clemens Nylandsted 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 SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. https://doi.org/10.1145/3025453.3025537

[79] Kathrin Probst, David Lindlbauer, Michael Haller, Bernhard Schwartz, and Andreas Schrempf. 2014. A Chair As Ubiquitous Input Device: Exploring Semaphoric Chair Gestures for Focused and Peripheral Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 4097–4106. https://doi.org/10.1145/2556288.2557051

[80] Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. HuddleLamp: Spatially- Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '14)*. 45–54. https://doi.org/10.1145/2669485.2676584

[81] 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 SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. 3913–3922. https://doi.org/10.1145/2702123.2702287

[82] Gonzalo Ramos, Kenneth Hinckley, Andy Wilson, and Raman Sarin. 2009. Synchronous Gestures in Multi-Display Environments. *Human-Computer Interaction* 24 (2009), 117–169.

[83] Jun Rekimoto. 1997. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '97)*. 31–39. https://doi.org/10.1145/263407.263505

[84] Jun Rekimoto, Yuji Ayatsuka, and Michimune Kohno. 2003. SyncTap: An Interaction Technique for Mobile Networking. In *International Conference on Mobile Human-Computer Interaction (MobileHCI '03)*. 104–115. https://doi.org/10.1007/

978-3-540-45233-1_9

[85] Jun Rekimoto and Masanori Saitoh. 1999. Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. 378–385. https://doi.org/10.1145/302979.303113

[86] Stephanie Santosa and Daniel Wigdor. 2013. A Field Study of Multi-Device Workflows in Distributed Workspaces. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp '13)*. 63–72. https://doi.org/10.1145/2493432.2493476

[87] Albrecht Schmidt. 2000. Implicit Human Computer Interaction Through Context. *Personal Technologies* 4, 2 (2000), 191–199. https://doi.org/10.1007/BF01324126

[88] Dominik Schmidt, Fadi Chehimi, Enrico Rukzio, and Hans Gellersen. 2010. PhoneTouch: A Technique for Direct Phone Interaction on Surfaces. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '10)*. 13–16. https://doi.org/10.1145/1866029.1866034

[89] Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A Cross-Device Interaction Style for Mobiles and Surfaces. In *Proceedings of the Designing Interactive Systems (DIS '12)*. 318–327. https://doi.org/10.1145/2317956.2318005

[90] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori. M. Inkpen. 2004. Territoriality in Collaborative Tabletop Workspaces. In *Proceedings of the Conference On Computer-Supported Cooperative Work (CSCW '04)*. 294–303. https://doi.org/10.1145/1031607.1031655 arXiv:arXiv:1011.1669v3

[91] Abigail J Sellen and Richard HR Harper. 2003. *The Myth of the Paperless Office.* MIT press, Cambridge, MA, USA.

[92] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. 2015. Gluey: Developing a Head-Worn Display Interface to Unify the Interaction Experience in Distributed Display Environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. 161–171. https://doi.org/10.1145/2785830.2785838

[93] John D Smith, Roel Vertegaal, and Changuk Sohn. 2005. ViewPointer: Lightweight Calibration-Free Eye Tracking for Ubiquitous Handsfree Deixis. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '05)*. 53–61. https://doi.org/10.1145/1095034.1095043

[94] Sony Business Solutions Corporation. 2018. Sony Digital Paper System DPT-RP1. https://www.sony.jp/digital-paper/products/DPT-RP1/

[95] Kentaro Takano, Hirohito Shibata, and Junko Ichino. 2014. Microscopic analysis of document handling while reading paper documents to improve digital reading device. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '14)*. 559–567. https://doi.org/10.1017/CBO9781107415324.004 arXiv:arXiv:1011.1669v3

[96] Kentaro Takano, Hirohito Shibata, and Kengo Omura. 2015. Effects of Paper on Cross-reference Reading for Multiple Documents: Comparison of Reading Performances and Processes between Paper and Computer Displays. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15)*. 497–505. https://doi.org/10.1145/2838739.2838745

[97] Ap Tarun, P Wang, and A Girouard. 2013. PaperTab: An Electronic Paper Computer with Multiple Large Flexible Electrophoretic Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '13)*. 3131–3134. https://doi.org/10.1145/2468356.2479628

[98] Jayson Turner, Andreas Bulling, Jason Alexander, and Hans Gellersen. 2014. Cross-Device Gaze-Supported Point-to-Point Content Transfer. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14)*. 19–26. https://doi.org/10.1145/2578153.2578155

[99] Jo Vermeulen, Steven Houben, and Nicolai Marquardt. 2016. Fluent Transitions Between Focused and Peripheral Interaction in Proxemic Interactions. In *Peripheral Interaction*. 137–163. https://doi.org/10.1007/978-3-319-29523-7

[100] Daniel Vogel and Ravin Balakrishnan. 2004. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '04)*, Vol. 6. 137–146. https://doi.org/10.1145/1029632.1029656

[101] Miaosen Wang, Sebastian Boring, and Saul Greenberg. 2012. Proxemic Peddler: A Public Advertising Display that Captures and Preserves the Attention of a Passerby. In *Proceedings of the International Symposium on Pervasive Displays (PerDis '12)*. 3–9. https://doi.org/10.1145/2307798.2307801

[102] Mark Weiser. 1991. The Computer for the 21st Century. *Scientific American* 265, 3 (1991), 94–104. https://doi.org/10.1145/329124.329126

[103] Andrew D. Wilson and Hrvoje Benko. 2010. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '10)*. 273–282. https://doi.org/10.1145/1866029.1866073

[104] Robert Xiao, Scott Hudson, and Chris Harrison. 2016. CapCam : Enabling Quick , Ad-Hoc , Position-Tracked Interactions Between Devices. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and*

Spaces (ISS '16). 169–178.

[105] Jishuo Yang and Daniel Wigdor. 2014. Panelrama: Enabling Easy Specification of Cross-Device Web Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. 2783–2792. https://doi.org/10.1145/2556288.2557199

[106] Dongwook Yoon, Ken Hinckley, Hrvoje Benko, François Guimbretière, Pourang Irani, Michel Pahud, and Marcel Gavriliu. 2015. Sensing Tablet Grasp + Micro-mobility for Active Reading. In *Proceedings of the ACM Symposium on User Interface Software & Technology (UIST '15)*. 477–487. https://doi.org/10.1145/2807442.2807510

[107] Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. 246–253. https://doi.org/10.1145/302979.303053

[108] Jamie Zigelbaum, Adam Kumpf, Alejandro Vazquez, and Hiroshi Ishii. 2008. Slurp: Tangibility, Spatiality, and an Eyedropper. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '08)*. 2565–2574. https://doi.org/10.1145/1358628.1358713

A PROTOCOL OF THE EVALUATION STUDY

Welcome to our study. During the study, we are going to show you some new and existing ways that people work with computers. I'm going to ask you some questions about these different techniques to learn how users such as yourself think about them and collect some feedback on them. Please remember you are here today to help us test the system, not you. Please relax and feel free to ask me any questions you might have. We want your honest opinion and feedback.

First, please review and sign the consent form. Please feel free to ask any questions you may have. Thank you.

Before the study, please complete this questionnaire about some demographic information. Your participant ID is X.

A.1 Classic Model

Thank you. First, let's start with how we use the computer today.

(1) Suppose that I am working on this Word document and I have a dataset here in the Excel application. If I realize that I made a mistake in the Excel spreadsheet, how can I make the changes?
(2) You can see that in this example, we move the mouse cursor to another application, and click, then we can start to do something on the new application. This is an example of what we call changing window focus: changing which part of the screen is able to receive input from the user. So, in this example, we say that we changed the Window Focus by moving and clicking the mouse.
(3) On Windows system, there is another way that we can do this. Using the keyboard, I can press and hold the "Alt" key and then press the "Tab" key. This will bring up a preview of all the things I have open and let me choose which application I want to input text to. Again, this is another example of changing the Window Focus, this time by pressing the keyboard.

A.2 Multi-Device Tracking

Now we are going to extend this concept of Window Focus from a single device to an entire environment that has multiple devices, including this laptop, desktop, keyboard, mouse, and these tablets on the desk. Using the cameras on the ceiling and the tapes on the devices, this system can track the location of all these devices. The tapes on the chair and on this cap also let the system determine where I am facing and where my head is pointed.

In this system, instead of only being able to give input to one device at a time, we can give input and focus to multiple devices. For example, the device we are touching, the device we are looking at, and the device our body is turning towards. We are now going to explore how touching, looking at, and turning towards different devices can change how we interact with computers.

A.3 Hand Focus

First, let's start with hand focus, which is quite common. As in the previous examples, when we use the mouse to click something, the mouse has our hand focus. Also, when you press the keys on the keyboard, the keyboard has our hand focus.

(1) The question is, I have so many devices here in my room. What does it mean if this device has my hand focus?
(2) In other words, when you pick up a device and touch it, what should the system do?

A.4 Gaze Focus

OK. That's how we deal with our hand focus. Next, let's talk about gaze focus. When we are looking at one device, we say that this device has gaze focus.

(1) (Gaze Copy) Here is an example of using gaze focus to help you copy application window between devices. Using gaze focus, we can copy the application to the target device, by pressing the title bar, looking at the target device, and releasing the mouse. So we can easily copy the application to the tablet, grab the device and read the document on the go.
    - What shall I do if I want to copy something from the tablet to the monitor?
    - Can you explain why does it happen?
(2) (Following Window) Now let's see another example using gaze focus. I am typing on this document, and now I decide to look at the reference document on another monitor. You can see that, even though I am still typing on this document here, I can see the characters I just typed on this monitor here.
    - Why might the text appear on the other monitors?
(3) Now you can try to move your head around. When you switch your gaze from the device A to device B, what does it mean for device A and B respectively?
(4) What other applications or scenarios do you think could also use gaze focus? Or you can just provide some comments for these two features, what do you like/dislike and how you would like to improve these techniques.
(5) As these little demos here have shown, by switching gaze or hand focus, we changed which device, and application, receives input from us.

A.5 Body Focus

We can take this idea even further. Because the system can track the movement of the chair, and thus your body, the system can also use this information to manipulate the device that has focus. If a device is in front of your body, we say that it has body focus.

(1) (Summary View) Here is an example using body focus: I am working on an article, and here are some support documents opened on the side. When I want to search something from them, sometimes I get confused because, at a glance, they look similar. But now, these devices show a summary of their content using a simplified view on the screen. I can now easily find what I want. When I grabbed this tablet back to me, its screen shows detailed information now.
    - Why do you think these devices show different types of content?
    - When will they change the form?
(2) (Gaze Search in local directory) Now let's see another example using body focus. I have some documents open on my desk, some are just in front of me with body focus, and others are far

away, outside my body focus. If I want to search for something, I can use the local search function, and see the result.
- Why might there be separate lists of devices on the search result page?

(3) Now you can sit in this chair and try to turn your body towards different devices. What do you want to do when you move one device into your body focus? What do you want to do when you move one device away from your body focus?
(4) What do you want to do when you sit on the office chair and rotate your body towards a group of devices, or turn away from them?
(5) Do you have any new ideas of using body focus? Or could you provide some comments for these examples I just showed to you?

A.6 Summary of User-Device Proxemics

Now we have shown three different ways to change which device is receiving input from us: using our hand, using our gaze, and using our body.

(1) Within the context of this system, what's does it mean when you are holding/touching something, looking at something, or turning towards something? Can you try to give us a rough description or definition?
  - When would you hold/touch something?
  - When would you look at something?
  - When would you turn towards something?

(2) I have shown you each of these individually, however we can also use them in combination. For example, in the local search function we just used: we make a selection by hand focus, and look at the target monitor with gaze focus, it showed some options, and we confirm with hand focus. In the result page, answers are sorted based on body focus.
(3) Another example is the following window application I showed you before: I am typing on this keyboard, and I looked at this tablet for some information, and this floating window displays here.
  - Can you try to analyze this example using the terms we just explored?

(4) Are there any other combinations of these three types of focus that would be useful? You can introduce your own types of focus, other than the hand/body/gaze stuff. Or you can talk about general comments about these examples using hand, gaze, and body focus.

(Break) Do you have any other questions about these three types of focus we just discussed? You can feel free to ask questions, and let's have a short break.

A.7 Inter-Device Proxemics

Now we are going to move onto a slightly different idea. As we mentioned before, every device in this system can be tracked and all the devices are connected to each other. So they can share information across all the devices. Previously, with body, hand, and gaze focus we were exploring how the user interacts with the devices, but now we are going to explore how the devices can interact with each other and the different types of relationships they can have with each other.

A.8 Aiming State

First, you probably have experience using a TV remote. Basically, we aim the remote to the TV, and we can control the TV with the buttons on the remote. Within this system, we can do the same thing. We call this the aiming relationship.

(1) (Remote Controller) Now, suppose that we want to control the printer. Instead of standing up and go to operate the printer, we can grab the tablet, and aim it to the target printer, then

we can see some control buttons on it, and the current status of the printer. Basically, our tablet has become a remote controller for the printer.

(2) (Universal Keyboard) Similarly, we can aim the keyboard towards other devices to control them.
   - Now, if I am working on this article, and I want to change something on this tablet, what shall I do?
   - I can rotate the keyboard towards it and the keyboard input is sent to the tablet. Can you explain the reason why the keyboard output is sent to the tablet?

(3) In what other applications or scenarios might it be helpful to aim one device at another?
(4) What other devices could be used in this aiming relationship? What do you think of these examples? Any ideas on how to improve them?

A.9 Broadcasting State

There are other relationships that we can create between devices as well. What happens if devices are not aimed at each other? Just like these tablets randomly placed on the desk. We say that they are broadcasting. Because every device is still tracked, they broadcast their information to other devices to let them know what is happening.

(1) (Device Radar) A few minutes ago I demonstrated how the keyboard could be aimed at a tablet. Before connecting the keyboard, we can see the black and white icons here, which tells us that the tablet has found a keyboard nearby and is ready for connection. This feature used the broadcasting relationship, that they know the existence of each other. Then we decide to connect them, and they are now aiming, and the icon turns blue.
(2) (Digital Labels) Here is another example. I have multiple documents open and I am reading them. I can filter them by clicking the labels, and then the other tablets with the same label will blink. We can also create new labels and apply to other devices, similar to how we use sticky notes.
   - Can you try to explain why these digital labels can be shared between them?

(3) What other applications might be helpful using the broadcasting state? You can start by thinking about what other information could be shared among broadcasting devices. What do you think of these techniques? Any ideas on improvement?

A.10 Expanding State

A third type of relationship that devices may have is called an expanding one. If I place the devices adjacently to each other, like this, they join together and form a connected surface.

(1) (Expanded Keyboard) Sometimes I have a document opened on my tablet, and I want to modify it. Suppose that I don’t have a physical keyboard around and can only use the built-in keyboard blocking part of the display, which is sometimes annoying. So, I grab another tablet, place it adjacently with it, then it serves as a keyboard. Since the two devices are adjacent to each other, they can connect to each other and the content can expand across them. The typed characters will be displayed on the target monitor.
(2) (Adjustable Reader) Let’s see another scenario. Now I am reading a paper on this tablet. I decide to view and compare two pages from the paper, so I grab another tablet, and put them together. It shows continuous pages of the document, and I can view multiple pages of the same document at the same time.
   - Can you try to explain why the second tablet shows the continuous page of the document, and why they can be scrolled together?

(3) What other information and features will be helpful when placing two tablets in the expanding state?
(4) What other devices could be placed together in the expanding state? What other applications might be helpful using the expanding state? When do you want to place devices together? What do you think of these techniques? Any ideas on improvement?

A.11 Piling State

We have one more relationship to introduce. Instead of being placed next to each other, devices can also be placed on top of each other, in a pile. This is just like how we sometimes organize paper documents on our desk.

(1) (Piling Hub) We often collect some data from multiple sources and put them in the same pile. The devices in the pile share information with each other and generate an aggregate view. By piling the devices on top of each other, I can create a figure on the top device, which will help me analyze the dataset.
(2) (Piling Portal) Now let's consider another situation, when we place multiple documents in a pile.
   - Can you try to guess what happens on the devices when I place them in a pile?
   - It shows a detail view of documents on the top device, and I can navigate all documents on this top most device. Can you try to explain why it shows the view of documents on the top device?
(3) What other information will be helpful on the top device? What other devices could be placed in a pile? What other applications might be helpful using the piling state? What do you think of these techniques? Any ideas on improvement?

A.12 Summary of User- and Inter-device Dimensions

We have just explored four different types of relationships that devices can have with each other: aiming, broadcasting, expanding, and piling. Do you have any questions about these four types of relationships?

The last thing I want to show you is that these relationships can also be combined with body, hand, and gaze focus that we explored in the first half of the study.

(1) (Enhanced Adjustable Reader) For example, we can have these multiple tablets connected on the desk, which are in the expanding state. When I look at one of the tablets, these devices will generate an integrated view with all notifications to the device I am looking at.
   - So, can you try to explain why this will work, using the terms we just talked about?
(2) (Piling Hub) Let's see another example. We have shown that we can create a figure for the dataset within the pile. After that, we can touch the figure, then look at the target monitor. Then the figure is copied to the destination.
   - Why does this work?
(3) Other than these examples, what other applications or scenarios might you want to use these aiming, expanding, piling states together with the hand/gaze/body focus? You can talk about the challenges you might have using computers in your daily life and think about whether these techniques might help. Or you can provide some comments on the advantage and disadvantage of this system.
(4) Other relationships that might be involved in this system?

So, in conclusion, we have explored different ways that this system can change which device you give input to: using not only the hand, but also the direction of your body and your gaze. We also explored that devices can have different relationships with each other, piled on top, adjacent to each other, pointed towards, or away and simply broadcasting.

Your feedback has been very valuable to our team and we appreciate your time. Do you have any questions for me? Thank you so much.

Received October 2018; revised December 2018; accepted February 2019