The Unadorned Desk: Exploiting the Physical Space around a Display as an Input Canvas

Doris Hausen
University of Munich (LMU)
Human-Computer-Interaction Group
Amalienstrasse 17
80333 Munich, Germany
doris.hausen@ifi.lmu.de

Sebastian Boring and Saul Greenberg
Department of Computer Science
University of Calgary
2500 University Drive NW
Calgary, Alberta Canada
[boring,saul]@ucalgary.ca

ABSTRACT
In everyday office work, people smoothly use the space on their physical desks to work with documents of interest, and to keep associated tools and materials nearby for easy use. In contrast, the limited screen space of computer displays imposes interface constraints. Associated material is either placed off-screen (i.e., temporarily hidden) and requires extra work to access (window switching, menu selection) or crowds and competes with the work area (e.g., as palettes and icons). This problem is worsened by the increasing popularity of small displays such as tablets and laptops. To mitigate this problem, we investigate how we can exploit an unadorned physical desk space as an additional input canvas. Our *Unadorned Desk* detects coarse *hovering over* and *touching of* areas on an otherwise standard physical desk, which is used as input to the desktop computer. Unlike other augmented desks, feedback is given on the computer's screen instead of on the desk itself. To better understand how people make use of this new input space, we conducted two user studies: (1) placing and retrieving application icons onto the desk, and (2) retrieving items from a predefined grid. We found that participants organize items in a grid for easier access, and are generally faster without affecting accuracy without on-screen feedback for few items, but were more accurate (though slower as they relied on feedback) for many items.

Author Keywords
Augmented desks, digital desks, peripheral interaction.

ACM Classification Keywords
H5.2 [Information interfaces and presentation]: User Interfaces: Input Devices and Strategies, Interaction Styles.

General Terms
Experimentation, Human Factors, Verification.

INTRODUCTION
In everyday office work, people naturally arrange documents, tools and other objects on their physical desk so that they are ready-to-hand, i.e., within easy reach and where they can be retrieved without actively searching for them. People are able to do so because they are aware of these objects' spatial location [15] and can coarsely acquire those that are in their peripheral vision. In contrast, working with computers requires almost everything to visually happen on-screen. Yet because space is limited, the so-called desktop metaphor usually separates object placement into one of several workspaces. First, the *primary workspace* holds the currently active document, which people normally work on. This space usually covers most of the screen or window. Next, the *secondary workspace* is the portion of on-screen space that contains a subset of artifacts related to the primary space's activities, e.g., icons and tool palettes.

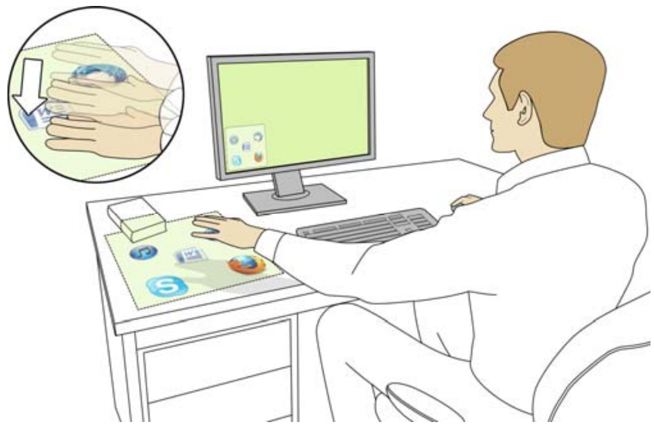

Figure 1. The *Unadorned* Desk: persons can interact with invisible, off-screen content on the desk. When hovering over the interaction area, feedback may be given on-screen. Touching an item (see callout) lets the person select that item.

Finally, the *off-screen workspace* holds the remaining artifacts, where users – through a series of operations – make them explicitly visible in a temporary fashion (e.g., menus, dialog boxes). Yet, there is a tension between these workspaces. The primary and secondary workspaces spatially trade-off: the primary workspace dominates screen space, which leaves less space for its surrounding artifacts. This is especially true for tablets and other devices with rather small displays. The secondary and off-screen workspace also trade-off: it is much easier to select items in the secondary space, but only a few can be held there. On the other hand, a huge number of items can be held in the off-screen workspace, but it is harder to select them (or to remember accelerator methods such as keyboard shortcuts) [20].

Instead of trying to fit everything on screen (either directly or through menus), we investigate using the unadorned (i.e., mostly unchanged) desk as a further space to contain artifacts. Our hypothesis is that people can then easily select commonly used functions (e.g., tools or other windows) lo-

Cite as: Hausen, D., Boring, S. and Greenberg, S. (2012) The Unadorned Desk: Exploiting the Physical Space around a Display as an Input Canvas. Technical Report 2012-1030-13, Department of Computer Science, University of Calgary, Calgary, AB Canada T2N 1N4.Includes accompanying video available at grouplab.cpsc.ucalgary.ca/papers/

cated on the desk's surface (see Figure 1) This has several advantages: (1) if we move artifacts from the secondary workspace to the desk, more display space can be allocated to the primary workspace. And (2), if we move artifacts from the off-screen workspace there, they will be easier to access. This also somewhat mimics the way we interact with everyday objects surrounding a document located on the desk (e.g., placing paints and brushes nearby for rapid retrieval while drawing).

We are particularly interested in using the desk *as is* with the smallest possible alterations. While detecting touches and hovering is easily possible with little augmentation using depth sensors such as LeapMotion[1] or Microsoft's Kinect camera, these technologies do not provide visual feedback. In previous work, feedback on regular desks was addressed through the use of projectors [19, 30], a tabletop computer as desk replacement [6], or by adding tablet computers next to the display as an interactive region [6]. In contrast, we consider an extreme stance, where we provide either no feedback or feedback on-screen and on demand. Both approaches keep desk instrumentation to a minimum, thus allowing for using *any* desk – such as at cafes – to serve as a workspace. Using computer science terminology, this is a lower bounds investigation: we want to understand to what extent interaction is possible using minimal or no augmentation (i.e., no visual targets or confirmatory feedback are provided on the desk).

To investigate how an *unadorned desk* can be used as an input space, we built a prototype comprising a Microsoft Kinect depth camera mounted atop a regular desk. Our Unadorned Desk tracks a person's hand and allows for *hovering over* and *touching of* content. As we were interested in how people can interact with off-screen content while keeping their attention on their main task, feedback is either not provided, or is given on-screen and on demand. We conducted two experiments: the first *memorability* experiment focused on how many virtual items participants could place on and later retrieve from the desk. In the second *acquisition* experiment, a varying numbers of virtual items were placed at predefined locations and participants had to retrieve them. Our work offers two contributions:

1. **A working prototype including several sample applications** that make use of an unadorned desk as an additional input space by augmenting the desk's surface with a depth camera.
2. **Experimental results that inform the design of such interactions** with respect to the amount of off-screen virtual items and the given on-screen feedback.

In the remainder of the paper, we revisit prior work related to our approach, describe our implementation and present a set of example applications. We outline the design of both of our studies, discuss results, findings and implications, and then list directions for future work. An accompanying video further illustrates the setup, the tasks, and the experiment.

[1] *LeapMotion*: http://leapmotion.com/

RELATED WORK

Our work builds on several areas of research that relate to how people organize documents on their desk, peripheral and bimanual interaction, interfaces without direct visual feedback, and augmented desks in general.

Organization of the Desk

We routinely and fluidly arrange and manage documents on our physical desks without focusing much attention on it. We can do so because the document's physical arrangement on the desk offers context information about the status and importance of certain tasks [8]. Malone studied desk organizations and found that files and (even more so) piles are the most commonly used arrangements on a desk [23]. Files are usually ordered systematically (e.g., in an alphabetic order). Piles, however, are not organized deliberately, and people thus more likely use spatial organization for retrieval. Associated tools and materials are generally arranged so they are available for reuse, such as by placing them nearby and ready-to-hand during active use, or by organizing them into known locations (such as desk drawers) [12].

Many systems try to bring this traditional way of organizing a desk into the digital world. In *Data Mountain* [27], people can organize browser bookmarks on a virtual table, which proved faster than bookmarking in Internet Explorer 4. *BumpTop* simulates the desktop by allowing users to arrange documents in a virtual 3D space using physics [1]. Customization features in graphical user interfaces let people spatially arrange tools around the graphical desktop [12]. In contrast to these systems, we are interested in using the desk *as is* instead of mimicking it on-screen.

Augmented and Interactive Desks

There is a history of work where digital content is brought onto the surface of the physical desk. This not only provides a workspace larger than the constraints of a computer display, but – in some systems – also allows both physical and digital artifacts to be used in tandem. Early work focused on (partially) digitizing the desk. The *Digital Desk* [30] uses a projected interface on the desk. A video camera senses interactions with fingers and/or a pen, and can capture content of paper materials (i.e. interacting with paper). Rekimoto et al.'s *Augmented Surfaces* [25] are projected extensions to a laptop's display on a table or a wall. Users are able to drag content from their laptop onto the table where it is visible all times. Thus, the table serves as visual extension to the laptop's display. Similarly, *Bonfire* [19] projects additional content next to a laptop's screen and allows touch input through cameras.

More recent prototypes augment the computer screen with a horizontal digital display ('surface') located underneath it. Surfaces typically allow for touch input, making sensing of user interaction easy (e.g., *Magic Desk* [6]). *Curve* [32] and *BendDesk* [29] merge both the horizontal desk area and the vertical display area into one gigantic high-resolution touch-sensitive display, where they seamlessly connected through a curve. Various studies investigated how particular touch regions on both the horizontal and vertical displays are used e.g., to show that the regions next to keyboard and mouse are best suitable for coarse interaction [6]. We build on this in that we use the areas left and right of the keyboard/mouse in our two studies.

Peripheral and Bimanual Interaction

Working with analogue documents on a desk often involves peripheral and bimanual interaction. *Peripheral interaction* offers coarse input styles in the periphery of the user's attention and thus quasi-parallel to the user's current primary task. The fundamental characteristics for peripheral interaction are human capabilities such as divided attention (i.e., processing two tasks in parallel without switching channels [31]), automatic and habitual processes (i.e., carried out with little mental effort and hardly any conscious control [3]), and proprioception (i.e., being aware of one's own body, its posture and orientation [7]). Today's prototypes incorporating peripheral interaction mainly rely on tangible interaction (e.g., [4, 11, 17]) or freehand gestures [16]. Our work adds to this in that we investigate how people interact coarsely in their periphery.

Bimanual (two-handed) interaction is the basis for peripheral interaction. While typically asymmetric, both hands influence each other leading to a kinematic chain [13]. Studies show that bimanual interaction can improve performance [9, 18]. At the same time, the body provides the kinesthetic reference frame, i.e., the user's sense of where one hand is relative to the body and the other hand [5]. Further, Balakrishnan et al. found that while separating visual feedback from the physical reference does affect performance, there is only a "remarkably small difference" when comparing interaction with and without visual feedback as long as "body-relative kinesthetic cues are available" [5]. We build on this as we separate feedback from interaction.

Interfaces without Direct Feedback

Spatial interaction does not necessarily need to rely on direct feedback or feedback at all. Gustafson et al.'s *Imaginary Interfaces* [14] make use of the visual short-term and visuospatial memory. By forming an "L" with the non-dominant hand the user creates a reference frame for spatial interaction. *Spin & Swing* [2] depends on an imaginary circle around the user. By turning themselves, users navigate through the content, which is displayed on a handheld device. The concept of *body-centric interactions* [10] takes it one step further by employing the space around a person's body to hold mobile phone functions. For example, *Virtual Shelves* [21] positions items in a hemisphere in front of the user's body. *Point upon Body* [22] uses the forearm as interaction area, which can be divided at most into six distinct areas. *GesturePad* [26] and *BodySpace* [28] use different body locations for different commands (e.g., tilting a mobile device near the ear allows for changing the current track). As with our system, no direct feedback is provided. These systems rely primarily on spatial awareness and kinesthetic memory. Due to proprioception, users have a good understanding of where items are located and can easily – even with closed eyes – place and retrieve such objects [24]. These findings inspired us to mimic regular desk use as means for interacting with digital content.

Allowing for input on the Desk

The *Unadorned Desk*, our working prototype, uses a Microsoft Kinect depth camera mounted on a tripod facing upside down (see Figure 2), so that it observes a sub-region of the desk within which a person could interact using the hand. The software supports arbitrary locations on the desk, thus allowing for either the left or the right side of the keyboard to be used as an input area. The prototype runs on an Intel i7 3.4 GHz computer to allow for fast processing (i.e., 640 × 480 pixel frames at 30 frames per second).

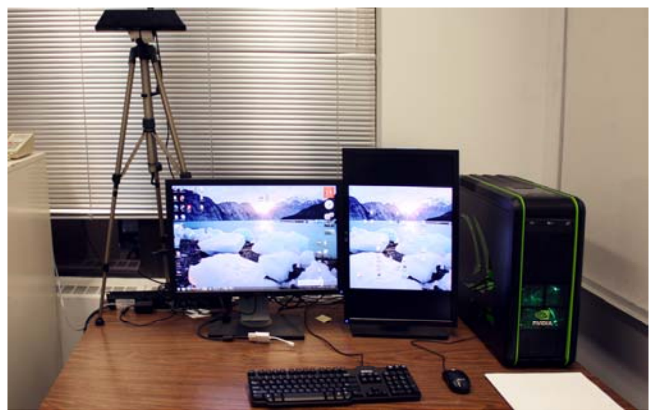

Figure 2. The Unadorned Desk: a standard desktop computer equipped with a Kinect to track a person's hand on the desk.

Detecting and Tracking a Person's Hand

We use the Kinect depth camera to gather hand information within the tracked region. The camera provides 16-bit depth images where each pixel in a depth frame encodes that pixel's distance to the camera in millimeters. At startup, the system takes a series of depth images (to reduce noise), averages them, and uses them as ground truth. Once running, it calculates the difference between the current depth frame and the calibrated depth image. The calculated difference image contains all points that are 'new' to the scene (e.g., a hand) with their distance to the desk. Using this point cloud, the system calculates the point of the hand closest to the corner of the interaction area that is the furthest away from the user (i.e., the tip of the middle finger). The vertical distance (depth) of that location to the desk further determines the hand's state: *touching* the desk if the value is lower than a given threshold, *hovering over* the desk if the value is higher, or *absent* if no hand is detected. On-screen feedback is optionally provided once the user's hand enters the interaction area. When the hand touches an item, the system performs the action associated with that item.

Example Applications

We built a set of example applications to demonstrate basic uses of off-screen space. The first application involves task-switching. When users bring their hand into the interaction area, the display shows all open windows (see Figure 3a). Users can then use their hand to select the window of interest (i.e., the hand's location is mapped to the windows' locations). The second application illustrates tool selection. People change tools in Adobe's Photoshop without moving the mouse to the icon of that respective tool (see Figure 3b) The third application (inspired by [17]) involves status changes, where people change their Skype status by tapping on the respective location on the desk (see Figure 3c)

Figure 3. Examples: switching between all windows (a), changing (b) Photoshop tools (b), or the Skype status (c).

Limitations

Our current implementation suffers from two limitations that restrict its deployment for everyday use. First, as with most optical tracking systems, the system is susceptible to false detections when sunlight hits the interaction area. Second, the system requires mounting a depth camera atop a desk, which is unsuitable for situations where rapid setup and teardown is required (e.g., temporary desks). This limits our ability to study the Unadorned Desk during anticipated everyday use. However, emerging technologies (such as LeapMotion) will likely overcome these limitations. Nevertheless, our prototype still allows us to evaluate how such interfaces can be used on unadorned desks.

EVALUATING OFF-SCREEN INTERACTIONS

In order to better understand how users can adapt to the novel input technology as well as how on-screen feedback for off-screen content would affect the interaction, we conducted two user studies. The first experiment aimed at understanding how people would spatially place onto the desk various content items that they would later retrieve. More precisely, we wanted to see whether people make use of special arrangements of their content. In the second experiment (which was tuned to use the results of the first study), we wanted to see how participants could accurately locate items placed in off-screen space as a function of the number of items in that space. The next section details the conditions and apparatus common to both experiments. Subsequently, we describe the unique conditions, tasks, participants, hypotheses and results of each experiment.

Conditions Common to Both Experiments

Although the tasks varied in both experiments, we had two conditions (used in addition to experiment-dependent ones) that were the same in both experiments: (1) the hand with which participants interacted in off-screen space, and (2) the type of feedback given during the task. In the following, we describe these two conditions in more detail.

Handedness: We chose to test our system with both hands. In the **dominant** hand condition, participants interacted with off-screen content using the hand they usually use to perform precise interactions (e.g., writing). In the **non-dominant** hand condition, they used the other hand. For each of the conditions, the interaction area was placed on the desk so that it was closest to the hand with which they had to interact in off-screen space (i.e., they did not have to reach left of the keyboard using their right hand). Thus we had two locations on the desk.

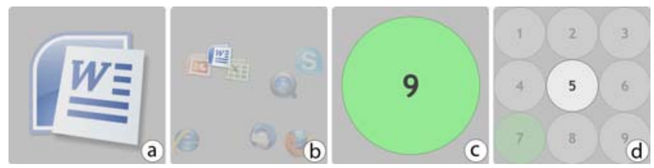

Figure 4. *First study:* a) Single feedback: Only the item closest to the participant's hand is shown. b) Full feedback: All items are shown with their correct spatial layout. Here the participant is currently hovering over "Word". *Second study:* c) Single: The target item is highlighted in green. d) Full: The participant's hand is currently hovering of item "5". *Both:* transparency encoded the distance to that item.

Feedback: We had three conditions for on-screen feedback. In the **No Feedback** (**None**) condition, participants did not receive any feedback on the computer's display, forcing them to rely solely on their spatial memory and proprioception. In the **Single Item Feedback** (**Single**) condition, participants only saw the item that was closest to their hand, with the distance being encoded through transparency. That is, as participants moved closer to a respective item, the item's icon became increasingly opaque (see Figure 4a,c). In the **Full Area Feedback** (**Full**) condition, participants saw all items in the interaction area with correct spatial layout. As in the *Single* condition, the transparency of items again changed based on the distance between them and the participants' hands (see Figure4b,d). That is, the item directly below the hand was more opaque than the surrounding items. The feedback area (400 × 400 pixels) was shown on-screen while a participant's hand was inside the interaction area and invisible otherwise. The feedback area was also located close to the interaction area (i.e., the bottom left or right corner of the display depending on handiness).

We used a within-subjects factorial design in both experiments: *2 Handedness* (*Dominant*, and *Non-Dominant*) × 3 *Feedback* (*None*, *Single*, and *Full*). The order of *Feedback* was counterbalanced across participants. To minimize changing the camera's location for *Handedness*, we alternated participants so that the first participant had all three

Feedback types with the *Dominant* hand and then again with the *Non-Dominant* one, while the second one started with *Non-Dominant* and continued with *Dominant*.

Apparatus and Setup

Both experiments used the previously described *Unadorned Desk*, with participants seated centrally in front of the computer's display. The depth camera captured a region of 40 cm × 36 cm (33.5 cm on the top edge due to slight camera distortion) next to the keyboard aligned with the desk's edges. For each *Handedness* condition, we moved the monitor, keyboard, mouse and chair to ensure that the participants are seated centrally in front of the display and close to the interaction area. The tracked region on the desk was empty. The computer display's background was set to a uniform color and had all desktop icons removed.

Hypotheses

We had similar hypotheses in both studies:

- H1. Participants' time to retrieve items would increase as the number of off-screen items increased.
- H2. Participants' error rate would increase as the number of off-screen items increased.
- H3. Participants' time to retrieve items would increase when no feedback was present.
- H4. Participants' accuracy would increase and thus their error rate decreases when feedback was present.

Participants

Each study used 12 participants. Sexes were mixed (first: study 6 female; second: 4 female), and ages ranged from 19 to 30 (average was 24). Each person only participated in one of the studies to minimize learning effects. Handedness varied, 9 were right-handed in the first study, and all in the second. Each session lasted up to 1.5 hours, and all participants received $20 as compensation for their time.

STUDY 1: PLACING AND RETRIEVING CONTENT

The purpose of our first study is to understand how participants would make use of the *Unadorned Desk* to organizationally place and later retrieve an item, and the effect of having an increasing number of items placed within that space. In particular, we were interested in (1) how they arrange a given number of items on their desk, (2) how they would memorize where they had placed items, and (3) how accurately they could retrieve them. Our *Handedness* × *Feedback* factorial design was extended to include a third *Sets* condition, which is the number of items participants had to place and retrieve in the off-screen space. To generate a more realistic scenario, we used icons of well-known, easily identifiable applications. These icons were: Word, Excel, Power Point, Firefox, Thunderbird, Skype, QuickTime, and Internet Explorer. For each condition, the amount of icons was ascending (to increase difficulty): 2, 4, 6, and 8 different icons had to be placed and then retrieved.

Tasks and Procedure

The experiment consisted of two phases for each combination of *Handedness* and *Feedback*: placing items and later retrieving them. We instructed participants to place icons off-screen in a position of their own liking. However, icons had to have a minimum distance of 52.5 millimeters (and 50 pixels respectively) to avoid overlaps of them, which would make retrieval more error-prone. Each set of icons they had to place was shown on the monitor during the placement task (see Figure 5a), so that participants were aware of all icons and could group them if that would aid their memory. To place an icon, participants first had to hit the spacebar to indicate they were starting the task, at which point timing began. Once the trial was active, they could move their hand into the interaction area and place the item by touching the desk's surface. When feedback was given, already placed icons were shown to give participants a feeling of the location of other items (as Figure 4a,b). Participants repeated this step until they had placed all icons in the current set in the physical off-screen space.

Figure 5. Commands shown during the study: a) Placing items. b) Retrieving items. c) Green border after pressing the spacebar indicating that the trial is active.

Once all icons were placed, participants had to retrieve them. Before a trial began, the system notified participants on-screen of which icon to retrieve (Figure 5b). They then had to hit the spacebar to activate the trial (Figure 5c). Participants would then retrieve that previously-placed off-screen icon. Retrieval worked exactly like the placement: hit spacebar for time measurement and touch a location to retrieve the icon. Afterwards, the system prompted them with the next icon until all icons were retrieved. If the wrong icon was retrieved, the participant was not informed, the trial was not repeated and the experiment continued. The error was recorded. For each *Feedback* and *Handedness* combination, participants placed 4 *Sets* of icons (*2*, *4*, *6*, and *8* icons) once and then retrieved each of them *4* times. We collected *24* placement sets (*480* icon retrievals).

For placement, we recorded all *x,y* locations (as the center) of placed icons. For the retrieval task, we measured the time from the beginning of a trial (i.e., hitting the spacebar) until they touched the desk's surface. We further recorded the location they touched, the distance to the actual icon (x,y location), and the amount of icons that were closer than the correct one (i.e., errors). We manually counted the participants' gazes, whether they looked at the interaction area, the feedback area, or both (the experimenter pressed a key for each gaze, which were recorded in an application). Finally, we asked participants to fill out a device assessment questionnaire: once after completing one *Feedback* and *Handedness* condition, and again at the end of our study.

Results

We used *heat maps* to qualitatively uncover how people would freely place different numbers of items on the desk. We then compared *retrieval time*, *retrieval accuracy*, and *gazes* to the interaction/feedback area during retrieval using separate repeated measures within-subjects analyses of variance (ANOVA). For pair-wise post hoc tests, we used Bonferroni-corrected confidence intervals to retain comparisons against $\alpha = 0.05$. When the assumption of sphericity was violated, we used Greenhouse-Geisser to correct the degrees of freedom. All unstated *p*-values are $p > 0.05$.

We performed a 2 × 3 × 4 (*Handedness* × *Feedback* × *Items*) within subjects ANOVA. As we did not find any significant main effects or interactions for *Handedness*, we aggregated over *Handedness* for all subjects in subsequent analyses. For heat map analysis, we mirrored interactions performed in the area right to the keyboard to bring those into the coordinate system of the one left to the keyboard.

Heat Map Analysis: Strategic Placement of Items

Item placement was not random. As shown in Figure 6, many participants tended to arrange items based on an imaginary grid. Further, participants followed other semantic patterns: First, some placed items in a single row as in the dock in Mac OS X. During retrieval with feedback, participants then hovered over that line to find the correct item. Second, some hierarchically grouped similar items together (e.g., all browser icons). They would later retrieve the icon by first going to the general group area containing that icon, and then selecting the particular icon. Finally, the more frequently they use an application based on their personal usage outside the study, the closer they would place it to the keyboard. Lesser used icons thus are further away from the primary interaction space. Participants did consider that areas further away would require more physical effort to access an item. However, all participants made use of the *entire* area, as they felt more comfortable to access items placed further apart from each other.

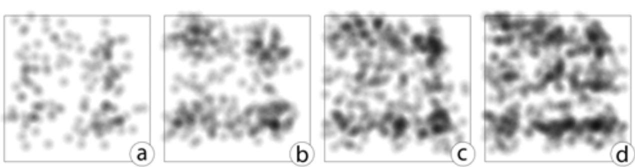

Figure 6. Heat maps ascending in item amount showing the placement for the interaction area located on the left of the keyboard (placements on the right side of the keyboard are included but mirrored). Users tend to arrange the items in grids.

We further calculated three *Distances* (*Closest*, *Average*, and *Highest*) between icons that they had placed in the off-screen space. Participants placed icons with an average item distances for all conditions between 207.4 and 231.6 millimeters (*M*=219.2; *SD*=9.7). To understand whether *Feedback* or the *Set* of icons had an influence on the distances between items, we performed separate 3 × 4 (*Feedback* × *Set*) ANOVAs for each *Distance*. For the closest distance, we found a significant main effect for *Set* ($F_{1.953,21.487}$ = 184.76, $p < 0.001$) and post hoc multiple means comparisons revealed that the distance increases with a decreasing *Set* of items (all pairs except *6* and *8* items differ with $p < 0.012$) regardless of *Feedback*. *Feedback* had an effect on the highest distance between icons, where we found significant main effects for both *Feedback* ($F_{2,22}$ = 15.49, $p < 0.001$) and *Set* ($F_{3,33}$ = 128.74, $p < 0.001$). Smaller *Sets* lead to lower distances between icons except for *6* and *8* items (all $p < 0.001$). More importantly, in the *None* feedback condition, participants placed items further away. The differences further increase with the *Set* size. Particularly for *8* items, *None* significantly differed from the other two (all $p < 0.045$), and from *Single* for *Set* sizes *2* and *6* (all $p < 0.046$). This shows that – when relying on feedback – participants felt more comfortable placing icons closer to each other. Interestingly, *Single* and *Full* did not differ for any *Set* size, and there was no significant difference between all three conditions for the *Set* with 4 items. This may be attributed to participants using the four corners of the area.

Retrieval Time

We compared retrieval times from the moment participants hit the spacebar until they retrieved the icon. We only took into account the retrieval times of those that were correct (even so, we did not find significant differences between retrieval times with and without errors). We performed a 3 × 4 (*Feedback* × *Set*) within subjects ANOVA and found significant main effects for *Feedback* ($F_{2,20}$ = 31.098, $p < 0.001$) and *Set* ($F_{1.609,17.698}$ = 15.583, $p < 0.012$). Figure 7a suggests that retrieval times slightly increase with larger *Sets*. More importantly, however, the *Feedback* influences retrieval times. Separate ANOVAs for each *Set* revealed that *No Feedback* was always faster than the other two *Feedback* conditions (all $p < 0.001$). Furthermore, the two conditions with visual feedback were more strongly affected by the *Set* of icons. Overall, *None* was the fastest (*M*=1.40s, *SD*=0.36s), followed by *Full* (*M*=2.47s, *SD*=0.88s), and *Single* (*M*=2.68s, *SD*=1.06s).

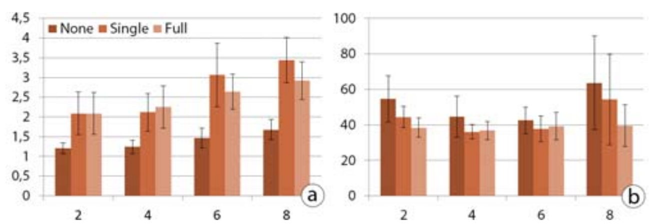

Figure 7. a) Retrieval time for one item for all feedback conditions and sets (in seconds). b) Distance for correct retrievals from the x,y coordinate of the corresponding item (in percent). (Error bars: 95% confidence interval.)

Accuracy of Retrievals

We first compared the accuracy (measured in distance between the touch point and the icon's center) of successful retrievals only (see Figure 7b). We performed a 3 × 4 (*Feedback* × *Set*) within subjects ANOVA and only found a significant main effect for *Feedback* ($F_{2,22}$ = 4.201, $p < 0.028$) but no effect for *Set* and no interactions. Overall, *Full* had the smallest offset between the touch and the icon's center (*M*=40.4 millimeters), followed by *Single* (*M*=45.3 millimeters) and *None* (*M*=54.0 millimeters).

We were also interested in the impact of *Feedback* on wrong retrievals (i.e., touch was closer to an incorrect item than to the correct one). We normalized the data (i.e., we divided the number of incorrect closer items by the maximum number of possible wrong items – 1 for a set of 2, 3 for a set of 4 etc.). We performed a 3 × 4 (*Feedback* × *Set*) within subjects ANOVA and found a significant main effect for *Feedback* ($F_{1.22,13.419} = 4.914$, $p < 0.039$). Post hoc tests, however, revealed that *None* was more error-prone than the other two conditions only for a *Set* with *6* icons ($p < 0.041$). We believe that – particularly in our case with no visual cues on the desk – participants made use of space to more accurately retrieve an icon. In summary, the chance for an erroneous selection with *None* is 20% (*SD*=12%), and 15% (*SD*=8%) for *Single* and *Full*.

Gaze Analysis

We told participants to minimize looking at the interaction area on the desk, and instead imagine that they were concentrating and looking at their primary on-screen task. We did not instruct them with respect to using the feedback window, that is, they could freely make use of it if available. We report gazes averaged across both placement and retrieval phase. Note that there were no gazes to the feedback area in the *No Feedback* condition.

For *Gazes* to the *Interaction Area*, we performed a 3 × 4 (*Feedback* × *Set*) within subjects ANOVA and found significant main effects for *Feedback* ($F_{1.126,12.383} = 7.948$, $p < 0.013$), *Set* ($F_{3,33} = 14.494$, $p < 0.001$) and a *Feedback* × *Set* interaction ($F_{2.15,23.645} = 8.618$, $p < 0.001$). Post hoc tests revealed that for *6* icons participants gazed at the interaction area more often in the *None* condition compared to *Single* ($p < 0.028$). For *8* icons, they gazed more often using *None* compared to the other two conditions (all $p < 0.017$). For *Gazes* to the *Feedback Area*, we did not test the *None* condition (as there was no feedback area) and performed a 2 × 4 (*Feedback* × *Set*) within subjects ANOVA and found a significant main effect for *Set* ($F_{2.121,23.334} = 7.274$, $p < 0.003$). Pairwise comparisons showed that *Gazes* to the *Feedback Area* increase with larger *Sets* (*2* and *4* differ from *6*, all $p < 0.045$, and *2* differs from *8*, $p < 0.022$). Overall, when *No Feedback* was presented, participants gazed at the interaction area on the desk more often (0.24 times per trial), compared to *Single* (0.11) and *Full* (0.12). In conditions that had *Feedback*, participants gazed at the feedback area 0.72 (*Full*) and 0.66 (*Single*) times per trial.

Discussion

During placement, we observed that participants used the whole interaction area, even though they stated that retrieval was easier if the item was placed closer to them. Placement was reasonably systematic, where each followed some kind of spatial organization. We also noticed an increased time for placement and found significant differences for icon distances in the *None* condition. From this, we believe that participants put more effort into finding a good arrangement (with reasonably spaced icons) to allow for an easier retrieval afterwards, which was especially important when there was no visual feedback of their placement.

During the retrieval stage, the condition without feedback caused two problems for participants: (1) they had to remember where they put items, and (2) they were not informed (or at least having a rough idea about it through the feedback area) whether they actually had correctly acquired an icon. Interestingly, participants stated afterwards that – when feedback was provided – they felt pressured to point accurately, which resulted in longer retrieval times. As one participant pointed out, he started to search instead of think, which slowed him down. Our analysis of gazes during such trials supports this view. Also, participants looked more at the interaction area when no feedback was given, although the amount they looked at it was less than the amount they looked at the feedback area for *Single* or *Full*. *Feedback* did help them to remember locations and thus increased their accuracy for larger *Sets*, but slowed them down.

Recall that these interaction techniques are to allow coarse interaction in the periphery (preferable with minimal attention). Our results suggest a suitable tradeoff between the item's sizes and the overall number of items. We observed that participants had problems memorizing their spatial layout with *6* or more items. Nevertheless, the results also indicate that participants were able to successfully retrieve *2* or *4* items – even without feedback. While the number of manageable items in real life scenarios could be quite large (e.g., participants may want to place many items meaningful to their task on the interaction area), others have argued that a small number of such items could comprise a large number of the actions people actually do [12]. Examples are frequently or recently used commands. Nevertheless, this first experiment suggests that having more items decreases accuracy. Quite possibly, our results could be affected by less than optimal placements of items on the desk, e.g., due to a lack of visual cues on the experiment's desk. For this reason, we conducted a 2nd study that spatially separated items into a grid (a preferred layout shown in this first study), and that did not require to memorize locations.

STUDY 2: TARGETING CONTENT

To prevent memorizing where exactly items were placed and eliminate the potential influence of unfavorable placement, we presented our participants a predefined layout, which was visible to them on the screen during each of the trials. Based on the findings in the 1st study, where participants had arranged items in a grid, we created grid-like layouts with pre-placed items. We added a variable *GridSize* with three levels: *2 × 2* (4 items), *3 × 3* (9 items), and *4 × 4* items (16 items). As we always filled the entire interaction area, item sizes differed depending on the *GridSize* (i.e., radius of 105, 70, and 50 millimeters). In this experiment, we were interested in getting more insights on item locations with respect to retrieval time, accuracy and errors.

Task and Procedure

Both – task and procedure – were similar to the retrieval task of the first study (though items are already pre-placed on the desk). At the beginning of each trial, the system showed participants which item they had to retrieve (see Figure 8). As before, they activated the trial by hitting the spacebar. Participants would then retrieve the respective item from off-screen space by touching the respective location. If they retrieved the correct item, the system prompted them with the next item to retrieve. If they touched the wrong one, the system notified them that the trial was incorrect, increased the item's error count, and asked them to retrieve it again. However, to avoid frustration, the system moved on to the next item after three failed attempts.

Figure 8. Commands instructing the participants which item they should retrieve (marked in green).

Participants had to retrieve all items in each *Grid Size* three times for each *Handedness* and *Feedback* combination, thus requiring every participant to perform 522 retrievals. However, the first block was excluded from the results as training block. We logged: *task time* from the moment the spacebar was hit until they either successfully retrieved the item or missed it; the *Euclidean distance* of the touch to the center of the item; and the number of *errors* (maximum = 3/item). As in the first study, we manually tracked if the participant looked at the interaction area on the desk, on the feedback area on the screen or both of them. After each *Feedback* and *Handedness* combination, participants filled out the same device assessment questionnaire used in the first study as well as a closing questionnaire in the end.

Results

We performed a 2 × 3 × 3 (*Handedness* × *Feedback* × *GridSize*) within subjects ANOVA. As in the first study, we did not find any significant main effects or interactions for *Handedness*. Thus, in subsequent analyses, we aggregated over *Handedness* across all participants. We also exclude all erroneous unsuccessful retrievals from our retrieval time and accuracy analyses, as we ended a trial if a participant was unsuccessful (three incorrect retrievals). Because of this, we excluded 6.5% of all trials.

Retrieval Time

Regarding retrieval time for an item, we performed a 3 × 3 (*Feedback* × *GridSize*) within subjects ANOVA and found significant main effects for *GridSize* ($F_{1.272,13,997} = 15.269$, $p < 0.001$) and *Feedback* ($F_{2,22} = 19.037$, $p < 0.001$). We further found a *GridSize* × *Feedback* ($F_{4,44} = 5.414$, $p < 0.001$) interaction. Post hoc multiple means comparisons revealed that for all *GridSizes* the retrieval time differed significantly for the *None* condition (in which participants needed less time) compared to the other two (all $p < 0.018$). In addition, for *Single* and *Full*, the retrieval time for the *4 × 4 GridSize* differed significantly from the shorter retrieval time for the *2 × 2* and *3 × 3 GridSizes* ($p < 0.001$). Overall, *None* was the fastest (*M*=1.68s), followed by *Single* (*M*=2.25s), and *Full* (*M*=2.33s). Figure 9a summarizes these results.

Accuracy of Retrievals

For the analysis of the accuracies of successful retrievals (measured in distance between the touch and the item's center), we had to normalize the distance since items in different *GridSizes* themselves had different sizes. We did so by dividing the offset by the maximum possible offset (i.e., the item's size). With the normalized data, we performed a 3 × 3 (*Feedback* × *GridSize*) within subjects ANOVA and found significant main effects for *GridSize* ($F_{2,22} = 39.318$, $p < 0.001$), and *Feedback* ($F_{2,22} = 4.918$, $p < 0.017$), but no *Feedback* × *GridSize* interaction. Pairwise comparison of different *GridSizes* across all *Feedback* conditions further revealed that participants were always more accurate for the smallest *GridSizes* ($p < 0.008$). For the 2 × 2 *GridSize*, participants were the most accurate with an offset of 46.9% of the item's width, followed by 3 × 3 (52.1%), and 4 × 4 (59.5%). Given the larger target areas in smaller *GridSizes*, however, participants were overall further away from the target's center (see Figure 9b).

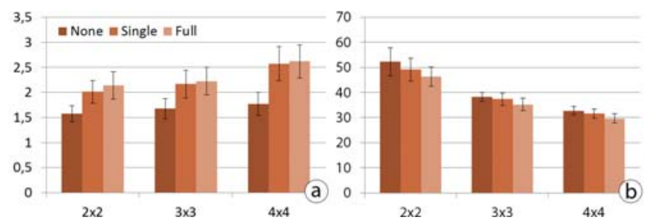

Figure 9. a) Retrieval time for one item for all Feedback conditions and GridSizes (in seconds). b) Distance for correct retrievals from the center coordinate of the corresponding item (in millimeters). (Error bars: 95% confidence interval.)

We normalized errors since we had a different amount of items depending on the *GridSize*. We divided the errors by the number of items in the grid for each trial. With these values, we performed a 3 × 3 (*Feedback* × *GridSize*) within subjects ANOVA and found significant main effects for *GridSize* ($F_{2,22} = 88.909$, $p < 0.001$), *Feedback* ($F_{1.309,14.4} = 10.587$, $p < 0.003$), and a *Feedback* × *GridSize* ($F_{2.126,23.385} = 4.036$, $p < 0.029$) interaction. Post hoc tests showed that the *None* condition differed significantly from the other two for the *2 × 2 GridSize* (all $p < 0.019$) and from the *Full* condition for the 4 × 4 *GridSize* ($p < 0.009$). However, *Feedback* conditions do not differ significantly for the *3 × 3 GridSize*. Overall, for all *GridSizes*, *None* was the most error prone (*M*=0.41, *SD*=0.23), followed by *Single* (*M*=0.22, *SD*=0.16) and *Full* (*M*=0.18, *SD*=0.10).

To understand the error-prone performance, Figure 10 is a heat map that visualizes the locations where participants had the most errors. As a general trend, one can observe that – for smaller amounts of items within the grid – the corner furthest away from the participant caused the most errors. However, the more items are placed in the grid, the more errors occur in the center, which can be explained by an easier targeting towards the borders of the grid due to a given reference frame (the edge of the desk). In a second analysis, we excluded the items further away: for *2 × 2*, we excluded the top left item, for *3 × 3* the three items furthest away, and for *4 × 4* the six items furthest away. We performed the same *3 × 3* (*Feedback* × *GridSize*) within subjects ANOVA using the reduced set and found significant main effects for *GridSize* ($F_{2,22} = 23.941$, $p < 0.001$), *Feedback* ($F_{1.332,14.648} = 9.973$, $p < 0.004$), but no *Feedback* × *GridSize* interaction. Post hoc tests revealed that both, *Single* and *Full*, differed significantly from *None* only for the largest *GridSize* (all $p < 0.046$). This supports our assumption that the corner furthest away is the most error-prone. However, *None* is still the least accurate across all *GridSizes*, with the best accuracy for *2 × 2* with 0.037 errors per trial (*Single*: 0.012, *Full*: 0.019).

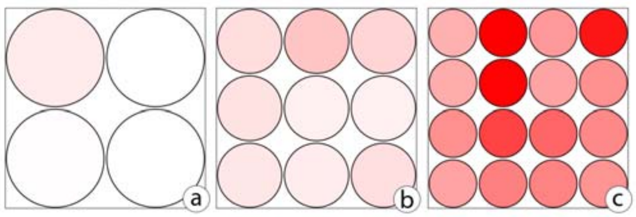

Figure 10. Heat maps showing errors (aggregated upon all feedback conditions, mirrored for the right interaction area) for GridSizes 2×2, 3×3 and 4×4. Saturation indicates errors.

Gaze Analysis
We instructed participants in the same way as we did in the first experiment. For *Gazes* to the *Interaction Area*, we performed a within subjects ANOVA on *Feedback* and found a significant main effect ($F_{1.136,12.495} = 10.485$, $p < 0.005$). Multiple means comparisons revealed that users gazed more often at the *Interaction Area* in the *None* condition compared to the other two (all $p < 0.023$). We again excluded the *None* condition for *Gazes* to the *Feedback Area*, and performed a within subjects ANOVA on the remaining two *Feedback* factors and did not find a significant effect.

Overall, *None* had the most gazes to the interaction area (0.23 times per trial), compared to *Single* (0.05) and *Full* (0.06). In *Feedback* conditions, participants gazed at the feedback area 0.69 (*Full*) and 0.65 (*Single*) times per trial.

Discussion
The second study re-enforces the findings from the first study. As before, *No Feedback* led to shortest retrieval times. Retrieval time also increased with a big grid (*4 × 4*) with feedback, which did not occur without feedback. As the number of items increased, size for an individual icon decreased. This forced participants to select more precisely for larger *GridSizes*. The offset we observed during *2 × 2* grids would not have been accurate enough for *4 × 4* sets. Users seemed to make use of space provided for smaller *GridSizes* but were able to adapt to bigger *GridSizes*.

No Feedback caused significantly more errors when the corner further away from the participant was included in the analysis. Similar to Magic Desk [6], where Bi et al. found that completion time was longer for areas further away from the keyboard, our participants had problems acquiring targets further away. When the items further away were excluded from analysis (i.e., only considering that half of the interaction area closer to the participant), the *No Feedback* condition only differed significantly from the other two for the largest *GridSize* (i.e., *4 × 4*). However, the error rate in the *4 × 4* grid was high regardless of the provision of feedback. We therefore assume that a grid with 16 elements is generally too large to be manageable in the periphery on an unadorned desk independent of feedback.

GENERAL DISCUSSION
As (H1) hypothesized, both studies showed that retrieval time increases as the number of items in the interaction area increases. While (H2) suggested that error rate increases as the number of off-screen items increased, this is only partially supported. We did not find a significant effect for more errors when increasing the item number (up to *8*) in the first study. Similarly, we did not find a significant effect in the second study for *3 × 3* (9 items), but we found a significant effect for *4 × 4* (16 items). Thus (H2) – more errors with more items – only seems to appear when there are at least 9 or more items. (H3) suggested that participants' time to retrieve items would increase when no feedback was present. Indeed, in both studies retrieval times were shorter when *No Feedback* was presented. Finally, (H4) suggested that participants' accuracy would increase and their error rate decrease when feedback was present. Yet our results at best show a tendency towards more errors and less accuracy with *No Feedback*. We did not find a significant effect for accuracy in the first study. There was a significant effect on errors for *6* items, but not for *8*. In the second study we found an effect for *2 × 2* and *4 × 4* but not for *3 × 3* comparing *No Feedback* to *Full Feedback* but not to *Single*. When only analyzing that half of the area closer to the participant, *No Feedback* only differs significantly from the other two for the largest *GridSize*. Our results therefore do not support (H4) and we may only report a tendency towards more errors and less accuracy with *No Feedback*.

The first study showed that participants made use of the whole interaction area, even with a rather small number of items. In the second study we found that items, located closer to the keyboard and mouse, are less error prone than locations further away. This suggests that a rectangular shape might not be the most suited interaction space. However, in in-situ experiments, this may change as participants would have better reference frames (i.e., items on the table that convey meaning) than just the blank, unadorned desk.

In general, our study showed that simple interaction on an unadorned desk is possible, albeit with a modest number of items. As the number of items was increased, both retrieval times and error rate increased as well. However, previous studies on peripheral interaction show that this interaction style needs to be trained and learned to be effective [4, 17]. This naturally is not possible in a short-term laboratory experiment. Abandoning feedback leads to faster retrieval times and functions (in terms of accuracy and errors) for a small numbers of items. Our findings suggest that the amount of items on the desk should be limited to less than ten. Similar to the shape of the interaction area, we expect this number to be higher if the desk contains more physical objects that serve as a visual cue or anchor.

Overall, participants enjoyed interacting with the unadorned desk, and considered it to be fairly easy. All were able to carry out the interaction equally well with their dominant and non-dominant hand, which strengthens our understanding that it is a peripheral interaction style. Interestingly, some of them were also irritated by this kind of interaction as they thought that the entire hand (and its palm respectively) acts as input, where in fact only a single point of the hand was tracked. Nevertheless, those participants adapted to the interaction fairly quickly.

CONCLUSION AND FUTURE WORK

In this paper, we presented the Unadorned Desk, which supports peripheral coarse interaction and extends the input- and workspace beyond a computer's display. The Unadorned Desk relies on hand tracking by a depth camera (Kinect). Our first studies showed that users are capable of interacting with virtual items on the desk, for small numbers of items even without additional feedback on the screen. It is a lower-bounds performance study, as we deliberately did not place anything on the desk's surface to indicate an item's virtual location.

There are still many unanswered questions for future work. Our first experiments were all carried out in an artificial lab setting, which brings with it usual concerns about external validity. The primary task was placement and selection, rather than doing one's actual work. The actual items had no special significance. Repeating the study in field cases could reveal nuances not seen in our lab study. Our interaction area was rectangular, of a given size, and uncluttered; all these could both be varied to see how it affects performance. It was also in 2D (albeit with a hover plane). Yet a 3D interaction space is possible, e.g., virtual piles where a user can navigate through it with the hovering hand. Finally, ours was a lower bounds study of an unadorned desk. There could be many possible ways of introducing modest adornments that indicate position. Although this would now

ACKNOWLEDGEMENTS

Funding provided by the NSERC / AITF/Smart Technologies Research Chair, and by NSERC.

REFERENCES

1. Agarawala, A., Balakrishnan, R. Keepin'it real: pushing the desktop metaphor with physics, piles and the pen. *CHI '06*.
2. Altakrouri, B., Kawsar, F., Kortuem, G. Spin&Swing: Spatial interaction with orientation aware devices. *Pervasive*, 2010.
3. Bakker, S., Hoven, E. van den, Eggen, B. Design for the periphery. *Eurohaptics*, (2010), 71-80.
4. Bakker, S., Hoven, E. van den, Eggen, B., Overbeeke, K. Exploring peripheral interaction design for primary school teachers. *TEI*, (2012), 245-252.
5. Balakrishnan, R., Hinckley, K. The role of kinesthetic reference frames in two-handed input performance. *UIST '99*.
6. Bi, X., Grossman, T., Matejka, J., Fitzmaurice, G. Magic desk: bringing multi-touch surfaces into desktop work. *CHI '11*.
7. Boff, K.R., Kaufman, L., Thomas, J.P. Handbook of perception and human performance. (1986).
8. Bondarenko, O., Ruud, J. Documents at hand : Learning from paper to improve digital technologies. *CHI*, (2005), 121-130.
9. Buxton, W., Myers, B.A. A study in two-handed input. *CHI*, (1986), 321-326.
10. Chen, X., Marquardt, N., Tang, A., Boring, S., Greenberg, S. Extending a mobile device's interaction space through body-centric interaction. *MobileHCI*, (2012).
11. Edge, D., Blackwell, A.F. Peripheral tangible interaction by analytic design. *TEI*, (2009), 69-76.
12. Greenberg, S. The computer user as toolsmith: The use, reuse, and organization of computer-based tools. (1993).
13. Guiard, Y. Asymmetric division on labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior 19*, 4 (1987), 486-517.
14. Gustafson, S., Bierwirth, D., Baudisch, P. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. *UIST*, (2010), 3-12.
15. Harrison, S., Dourish, P. Re-place-ing space: The role of place and space in collaborative systems, *CSCW*, (1996), 67-76.
16. Hausen, D., Boring, S., Polleti, J., Butz, A. Exploring design and combination of ambient information and peripheral interaction. *DIS Work in Progress*, (2012).
17. Hausen, D., Boring, S., Lueling, C., Rodestock, S., Butz, A. StaTube: Facilitating state management in Instant Messaging Systems. *TEI*, (2012), 283-290.
18. Kabbash, P., Buxton, W., Sellen, A. Two-handed input in a compound task. *CHI*, (1994), 417-423.
19. Kane, S.K., Avrahami, D., Wobbrock, J.O., Harrison, B., Rea, A.D., Philipose, M., LaMarca, A. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. *UIST*, (2009), 129-138.
20. Lane, D.M., Napier, H.A., Peres, S.C., Sandor, A. Hidden costs of graphical user interfaces: Failure to make the transition from menus and icon toolbars to keyboard shortcuts. *International Journal of HCI 18*, 2 (2005), 133-144.
21. Li, F.C.Y., Dearman, D., Truong, K. N. Virtual shelves: interactions with orientation aware devices. *UIST*, (2009), 125-128.
22. Lin, S., Su, C., Cheng, K., Liang, R., Kuo, T., Chen B.-Y. PUB-Point upon body: Exploring eyes-free interaction and methods on an arm. *UIST*, (2011), 481-487.
23. Malone, T.W. How do people organize their desks?: Implications for the design of office information systems. *ACM Transactions on Office Information Systems. 1, 1* (1983).
24. Mine, M.R., Brooks Jr., F.P., Sequin, C.H. Moving objects in space : Exploiting proprioception in virtual-environment interaction. *SIGGRAPH*, (1997), 19-26.
25. Rekimoto, J., Saitoh, M. Augmented surfaces: a spatially continuous work space for hybrid computing environments. *CHI*, (1999), 378-385.
26. Rekimoto, J. GestureWrist and GesturePad: unobtrusive wearable interaction devices. *Wearable Computers*, (2001), 21-27.

27. Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D., Dantzich, M.V. Data Mountain : Using spatial memory for document management. *UIST*, (1998), 153-162.
28. Strachan, S., Murray-Smith, R., O'Modhrain, S. BodySpace: inferring body pose for natural control of a music player. *CHI EA*, (2007), 2001-2006.
29. Weiss, M., Voelker, S., Borchers, J. Benddesk: Seamless integration of horizontal and vertical multi-touch surfaces in desk environments. *Adjunct Proceedings ITS*, (2009).
30. Wellner, P. Interacting with paper on the DigitalDesk. *Communications of the ACM 36*, 7 (1993), 87-96.
31. Wickens, C.D., McCarley J.S. Applied Attention Theory.
32. Wimmer, R., Hennecke, F., Schulz, F., Boring, S., Butz, A., Hussmann, H. Curve: revisiting the digital desk. *NordiCHI*, (2010), 561-570.