A Comparison of the Affordances of a Digital Desk and Tablet for Architectural Image Tasks

Ame Elliott and Marti A. Hearst
102 South Hall
University of California, Berkeley
Berkeley, CA 94170
{ame,hearst}@sims.berkley.edu

The physical context of architectural design includes large workspaces, typically drafting tables covered with piles of images and sketches. We are investigating if and how a large computerized workspace can be integrated usefully into such a design environment. To this end, we compared a large computerized desktop (digital desk) to a standard desktop computer and a small tablet environment for two typical architecture design tasks: sketching and image sorting. For the sketching task, the participants' preferences were evenly divided between the digital desk and the tablet. For the image sorting task, the desk was the least preferred environment, and produced significantly higher sorting times and more mistakes. Investigation into the causes of this difference yielded several interesting findings, including: the height of the participant was significantly associated with their speed on the sorting task, the larger image size available on the desk compensated for its poorer resolution in subjective preferences, and the quality of the alignment of the pen was an important factor both for preference and scoring results in the sketching task. Highly responsive pen input devices seem critical for user satisfaction not only for sketching, but also for image sorting; the effects of large display spaces are difficult to isolate from the limitations of input device. This paper elaborates on these findings and considers the implications for the design of user interfaces for image manipulation, in particular interaction techniques appropriate to using pen-input with large display surfaces.

1. Introduction

The physical context of architectural design includes large workspaces, typically drafting tables covered with piles of images and sketches. We are investigating if and how a large computerized workspace can be integrated usefully into such a design environment. Current computer support for architectural design focuses primarily on the last stages of the design process, that is, in aiding the production of drawings describing fully-designed buildings (Kalay and Carrara, 1994). Previous studies of architectural practice explore the organizational and economic aspects of running an office without addressing architects' activities during the design process (Larson, Leon, and Bolick, 1983; Blau, 1984; Cuff, 1991; Stevens, 1998). The early phase of design in architectural work is generally not well-understood, but recent results show that the process is creative, open-ended, and dynamic, consisting of retrieving and looking at images of precedents and making sketches (Elliott, 2000). Making sketches while looking at an image is one way architects determine the relevance of that particular image to a task at hand. Sorting images into groups is another way architects gather design information from images. The categorization process helps clarify design intent and allows relevant images to emerge via their relationship to members of a group. A recent overview of visual databases in architecture reiterates the importance of looking at images and sketching, but falls short of attempts to explore how these databases could be accessed in the context of the professional practice of design (Koutamanis, Timmermans, and Vermeulen, 1995).

Our ultimate goal is to devise computerized tools to better support architects in their work, particularly in their use of images. Before evaluating new image search interfaces in such an environment, we set out to determine if a large desk-like computerized display is useful for simpler tasks. To help determine this, we conducted a user study comparing the use of three different environments:

- A digital desk, a 35"x 26" display housed in a table the size of a drafting table using pen-based input
- a tablet, an 8.5" x 11" LCD display with pen-based input; and
- a standard monitor with mouse-based input.

Several physical properties of the digital desk suggest it would be better than the standard keyboard-mouse-monitor environment for the sketching and image browsing tasks common to the early phases of architectural design. The desk has a pen input device, which allows architects to take notes, select objects, and sketch in a way that should be natural and comfortable. The large display size lets users work with many images simultaneously, and the upright orientation of the display allows architects' to work standing, or seated on a high stool, just as they often work in non-computerized environments.

Numerous research projects use large displays as part of an attempt to support non-computerized work practices in electronic environments, including the DigitalDesk (Newman and Wellner, 1992), Tivoli (Pedersen et al., 1993), InfoWall (Rekimoto and Saitoh, 1999), HoloWall (Rekimoto and Matsushita, 1997), metaDESK (Ullmer and Ishii, 1997), i-LAND (Streitz et al., 1999), Flatland (Mynatt et al., 1999), and the Interaction Mural (Guimbretiere, Stone, and Winograd, 2001). The advantages of using a large display for direct manipulation are well documented. Fitzmaurice and Buxton (1997) advocate using specialized input devices for specialized tasks, suggesting that a configuration designed for architects' image manipulation needs would be more effective than a one-size-fits all setup of mouse and keyboard. These systems are intended to serve as either relatively static artifacts of handwritten notes, or as dynamic capturing media designed to support collaborative work. Our intention is to support tasks relating to browsing, searching, and using large collections of information.

In this paper we report the results of a comparative usability study of architects performing two tasks: image sorting and sketching, in each of the three different environments. The results of the user study disprove our initial hypothesis that the desk environment would be unilaterally more helpful for architectural tasks than the tablet or the monitor. Instead, for qualitative measures, the desk did only as well as the tablet for the sketching task and much worse than either the tablet or the monitor for image sorting. Quantitative results show more sorting mistakes on the desk than other environments and significantly longer times to complete the image sorting task on the desk. Interestingly, we found that the height of the participant was significantly associated with their speed on the sorting task; shorter participants were significantly slower on the desk, but not so on the other environments. The monitor and mouse were preferred over both devices for the sorting task. This may be the result of a combination of factors: the sorting application is designed to be used with the pointing acuity of the mouse, and the monitor screen size strikes an acceptable balance between having images large enough to view easily but not so large as to require large head movements to see all the images.

Another interesting result was that the desk's poorer screen resolution did not register negatively with the participants. For subjective measures, the larger image size apparently made up for the relatively poorer resolution.

Finally, the quality of the alignment of the pen was an important factor both for preference and scoring results in the sketching task. The alignment of the pen to the desk is worse than that of the pen to the tablet, due in part to parallax errors increasing with display size. Architects who hold the pen at an angle while drawing are particularly hampered by the alignment problem.

The remainder of this paper, after discussing related work, describes the user study methodology, analyzes the results in more detail, highlights relevant related work, and outlines the implications of the findings for the design of user interfaces for image retrieval on large displays.

2. Methodology

2.1 Overview

The study consisted of two tasks: sorting sets of images and sketching a copy of a line drawing. These tasks were designed to be more constrained, and thus more easily evaluated, than the open-ended tasks common to the early phases of the design process. We assumed that if one tool dominated the others in these tasks, it would then make sense to evaluate it further with more open-ended tasks. The sorting problem was a timed task in which participants were encouraged to arrange images into a pre-determined rectangular layout as quickly as possible. The sketching task was an untimed task in which participants were asked to copy an image from a physical piece of paper. Three different computer environments were used in a within-subjects design: a desktop PC with a mouse, a desktop PC with an LCD tablet and stylus, and a digital desk with a stylus.

2.1.1 Equipment. The hardware specifications of each of three devices appear in Table 1. For the image sorting task, participants used the slide sorter view of

	Monitor	Tablet	Desk
			
Input Device	Logitech 3-button mouse	Pen type 1	Pen Type 2
Display	Monitor	LCD	Rear-Projected
Display Size	11" x 12" (43 x 30cm)	11" x 8.5" (28 x 21.6 cm)	31" x 26" (79 x 66 cm)
Resolution	1280 x 1024	1024 x 768	1280 x 1024
Pixels/Inch^2	6425	8411	1626
Images/Inch^2	.17 images	.37 images	.04 images
Image Size	2" x 1.4"	1.43" x 1.06"	4" x 2.9"

Figure 1
Comparison of devices used in the study.

Microsoft PowerPoint to order the images. Each display showed the same number of images at a time (5 rows and 7 columns), and every set contained exactly 35 images.[1]

For the sketching task, participants used Adobe Photoshop with the drawing window open to full screen size. The cursor was set to draw as a thick black pen with width equivalent to the line weight of the example sketch. No tool palettes or other windows were visible, and participants were instructed to react as though they were given a permanent black magic marker and a piece of blank paper. They could start over as many times as they wanted, but could not erase any marks.

Participants were given a workbook containing informed consent paperwork, surveys, written versions of the oral instructions for the tutorial on how to use the equipment, examples of the sketches to be copied, written versions of the oral instructions for how to sort the images, and examples of correctly sorted sets of images. Participants controlled their own workbooks throughout the study and could refer to them at any point. During each sorting task participants were given a large template for easy reference showing the correct order for the images as a supplement to the workbook.

1 We increased the double-click speed to maximum in all conditions to avoid the undesirable effect that occurs when a user inadvertently double-clicks an image (it changes the display instead of selecting an image). During the tutorial participants were shown this possibility and instructed about what to do if it should occur.

Device Order	Number of Participants		
	7	7	7
First	Tablet	Desk	Monitor
Second	Monitor	Tablet	Desk
Third	Desk	Monitor	Tablet

Table 1
The experiment design.

2.2 Experiment Design

Twenty-one paid participants, 13 male and 8 female graduate students in architecture, completed a within-subjects study. The procedure consisted of completing informed consent paperwork, a background questionnaire, tutorial and ergonomic adjustment on each device, the sketching task, a written survey and oral follow-up questions about the sketching and image sorting tasks. The participants' interactions with the systems were recorded on videotape and with still photos. The participants spent an hour and a half working one-on-one with an experimenter.

Ergonomic adjustment on the digital desk and tablet included offering the participant three chairs of different heights and the option of standing, adjusting the angle and height of the display surface so that the participant understood the range of possible positions, and demonstrating two lighting options to reduce glare and eye-strain. The adjustments for the monitor included only two chair options and lighting options. The device order, derived from a Latin square design, was the same for both sketching and sorting tasks. Table 1 below shows the experiment design.

For the sketching tasks participants spent a few minutes getting familiar with the environment and using the input device for sketching. After completing two trial runs that were not used in analysis, participants sketched a set of six abstract figures. The participants then moved to the other two devices to complete additional trial runs and sketch different groupings of abstract figures. The motivation for sketching abstract figures is to limit the semantic implications of copying a sketch of a famous building or landscape and isolating aspects of the mechanics of sketching. Isolating the architects' responses to the mechanics of each of the pens is important because little is known about how the mechanics of pen input for professional design tasks work; the bulk of the work with pen computing is as a general purpose input device, e.g., (Greenstein and Arnaut, 1988; Kurtenbach and Buxton, 1991; MacKenzie, Sellen, and Buxton, 1991; Accot and Zhai, 1999).

For the image sorting tasks, participants completed a tutorial consisting of sorting numbers into ascending order (to allow them to become familiar with the slide sorter view of PowerPoint for each display and input device). After completing the tutorial, the participant sorted three sets of images in each environment. Each consecutive image set increased in complexity. The first image set consisted of colorful shapes that the participants were asked to group into a pre-determined order (for example, red triangles first, followed by yellow circles). The second set of images consisted of sorting a mixed group of photographs by placing people

Figure 2
View of the starting condition of mixed photographs before the participant began sorting.

first, then animals, and then objects. See Figure 2 for an example.[2] The third and final set of images is the most architectural, consisting of interior and exterior photographs of buildings. Participants were asked to place images of the built environment into three groups based on the architectural concept of scale. Scale was based on how much of the environment an image contained: part of one building or room, one entire building, or a part of a city or landscape with space for many buildings.

Each of three tasks involved arranging the images into groups with similar images. In all cases the desired order was explained verbally and by an example page of images in the correct order posted nearby. For example, a printed page showed how the correctly sorted images would look with people first, then animals, and finally things. Participants were instructed to use only those three categories and not to worry about, e.g., placing vehicles with vehicles or fruit with fruit. They were also instructed to make quick decisions and if they felt an image's category membership was ambiguous, to make their best decision quickly and to move on.

Each participant completed a unique but equivalent version, randomly deter-

2 The images are available for free use and distribution from Microsoft's ClipArt web site, http://cgl.microsoft.com/clipgallerylive.

	Desk	Tablet	Monitor
Sketching Preference	10	10	1
Sorting Preference	3	7	11

Table 2
Participants' preferred environment for sketching and sorting.

mined, of each kind of image sorting task. For example, a participant would sort shape set A1, photo set B1, and building set C1 on the desk, shape set A2, photo set B2, and building set C2 on the tablet, and so on.

2.3 Evaluation Measures

The data gathered during the test included three written surveys using five-point Likert scales, the sketches, the number of mistakes made in the image sorting task, the time taken to complete image sorting (both total time and time of selected individual transactions), and transcriptions of responses to open-ended verbal questions about device preferences. The first survey recorded the participants' background in architecture, education, work habits, and computer experience. The other two surveys consisted of comparative questions about each of three devices for the sketching and image sorting tasks. Additionally, the sketches were evaluated by a panel of five designers. This procedure is described in detail below.

3. Results

In this section we first summarizes the participants' device preferences for both tasks. We then report both quantitative and qualitative results for the image sorting task, followed by qualitative and quantitative results for the sketching task.

3.1 Subjective Preferences

The participants had widely varying preferences among the three environments. In the sketching task, preferences were evenly split between the desk and the tablet. Ten participants preferred the desk, 10 the tablet, and 1 the monitor and mouse. For the image sorting task, 3 preferred the desk, 7 the tablet, and 11 the monitor and mouse. All of the participants who preferred the desk for sorting also preferred it for sketching. Despite the variety of the participants' backgrounds, no correlation were found between any device and the following individual characteristics: years of architecture or computer experience, frequency of free-hand drawing, or most common work environment.

Although the participants' backgrounds did not correlate to device preference, we did discover unexpectedly that height correlated to device preference for image sorting. A review of videotapes showed shorter people struggling to reach the edges of the desk, despite the ergonomic adjustment part of the study. In our initial evaluation gender was weakly correlated to device preference (women dispreferred the desk). However, once two shorter men were added and one taller woman removed from the group, the height correlation became highly statistically significant ($p <= 0.01$, computed as a single factor analysis of variance

(ANOVA)).[3] The threshold for dispreferring the desk is height less than 5'6" (167 cm). Of the participants, 9 were shorter than 5'6" and 11 were taller. The average height for female participants was 5'3," the same as the United States average.[4] We estimate the average height for men in the study to be 5'7", below the national average of 5'9."

Participants' height is usually not measured in the course of a user interface user study, and few references to its implications for user interface design are available, other than oblique references in accessibility issues for the wheelchair bound. One of the few to mention height and arm length is Kroemer (1988) writing about the ergonomics of VDT design. These observations go little further than indicating that elbow height and forearm length have implications for input devices and writing, although the article does mention some preferred angles.

The hypothesis that height contributes to device preference is supported by the finding that all three participants who preferred the desk for image sorting were 5'6" or taller, and they all chose to stand rather than use any of the three seating options. This suggests that the quick visual and physical access to all parts of the screen is an important part of device preference for sorting. However, device preference for sketching did not correlate with any measure, including height and seating preference.

3.2 Sorting Task: Quantitative Results

Two quantitative measures distinguished the desk from the other devices in the sorting task: number of mistakes and task completion time. Participants made more sorting mistakes on the desk than on either of the other devices. Only icons and mixed image sets were used in the analysis of mistakes, thus there were $21 * 2 = 42$ different sorting instances evaluated for each device.[5] The frequency of errors for the icon and mixed image sets was 17 on the desk (40% of the sorting sessions on average), 11 on the tablet (26%), and 9 on the monitor (21%). The distribution of the mistakes on the screen is similar across devices, as shown on the following page in Figure 3.

Mistakes occurred predominantly in the periphery of all three devices, but no single participant produced more than 3 errors. The average was 0.8 mistakes/participant on the desk, 0.5 on the tablet, and 0.4 on the monitor. Given the higher number of mistakes on the desk and preference for either the tablet or monitor for sorting, the lack of correlation between device preference and number of mistakes is surprising. There is no statistically significant correlation between device preference and number of mistakes ($p = 0.3$, computed as a two-variable

3 The height of one male participant was not recorded because we were unable to contact him after the study. Since height was not something we were initially evaluating, our background survey did not include this information, and we collected it after making observations from the videotapes.
4 According to the Center for Disease Control (CDC) National Center for Health Statistics, as of February 2000. (http://www.cdc.gov/nchs/fastats/bodymeas.htm).
5 The building sorting task was too subjective to mark any mistakes, and the mixed photographs were only marked as a mistake if something was patently wrong. For example, a photo of a man on a donkey could be correctly sorted with people or with animals, but placing a photo of a plane in between photos of birds would be a mistake; the plane could only belong with things.

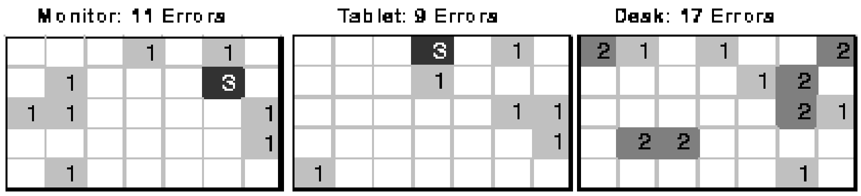

Figure 3
Position on display surface of images left incorrectly sorted and the total number of errors for each location.

regression).

Also surprising is that despite the obvious trouble that shorter participants had reaching the top row of the desk and the high number of mistakes occurring in that region, there is no statistically significant correlation between height and the number of mistakes ($p = 0.48$). Instead, the difficulty reaching the edges of the desk contributes to the second quantitative difference in desk performance: slower speed sorting images on the desk. Participants with height less than 5’6” were slower than the others at image sorting on the desk ($p <= 0.01$, ANOVA), even though the two group’s speeds are not statistically different on the tablet. The difference between the two group’s speeds is most pronounced on the building sorting task. The average speed sorting building images on the desk for participants 5’6” and above was 92 seconds, whereas for participants below 5’6” it was 158 seconds, 172% greater. Figures 4 and 5 show the times to complete the image sorting tasks for participants above and below 5’6”.

Because building sorting is the most complex task, we would expect user behavior for this task to be different than the others. This greater complexity is a deliberate attempt to simulate the open-ended image organization tasks architects typically perform while searching. Close analysis of the individual images and a constant re-thinking of the categories is necessary to determine if a photo of part of a high-rise building should go in a group with a photo of the skyline of a city, or with a picture of the front porch of a small cottage. This kind of interaction is a better approximation of architects’ work practices than sorting images into pre-defined categories.

To further investigate the timing differences, we used the videotape to analyze the behavior of three representative participants (fast, medium, and slow) completing the shape sorting task on the tablet and the desk. We separated out the time taken for actually moving the slide (what we call transaction time) from the rest of the time taken in the session (what we call selection time).

To calculate average transaction times per user on each device for a fixed distance, we recorded only ”clean” transactions – position changes in which the user selects one slide and puts it into the intended place on the first try without hesitation. This division of selection time and transaction time is analogous to

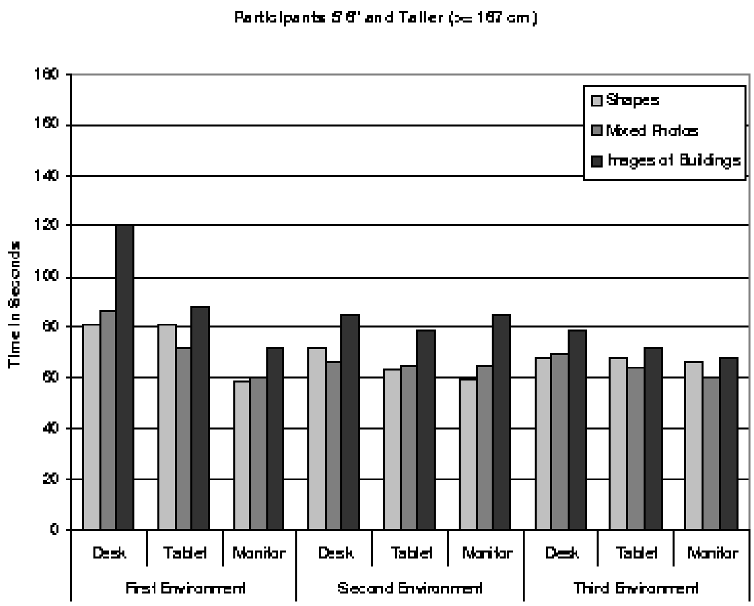

Figure 4
Image sorting times for participants 5'6" and taller. The first, second, and third environment refers to the order in which the participants used the devices, respectively.

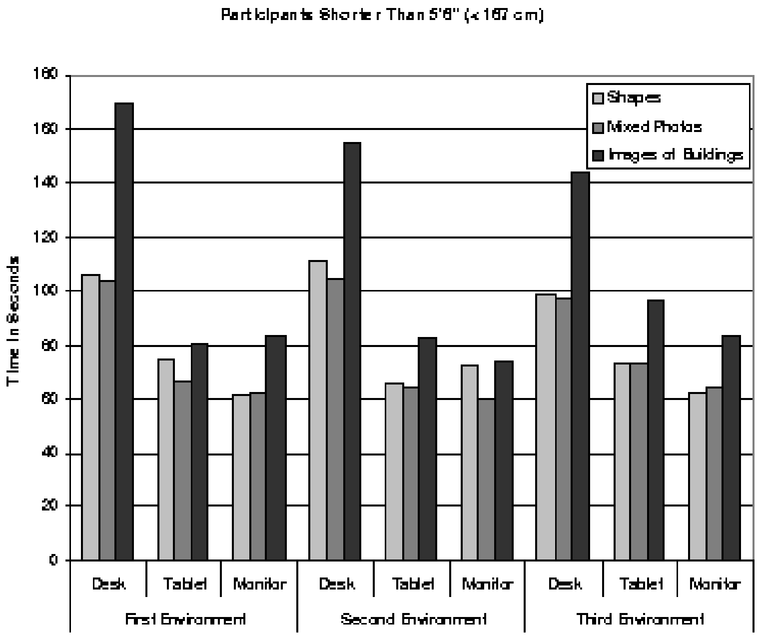

Figure 5
Image sorting times on the desk for participants less than 5'6" are significantly higher than for the taller participants for the building sorting task.

MacKenzie et al's separation of the point-select and drag-select task (MacKenzie, Sellen, and Buxton, 1991). Because these data report on times for the shape sorting task, it is possible to infer intended position because the shapes have a correct and incorrect location in the final arrangement. Looking only at clean transaction times, the participants were on average 131 milliseconds slower on the desk than the tablet.

Despite the slight differences in average sorting time for clean transactions, the differences in the total time needed to complete the shape sorting task are substantial: the slow participant took 99 seconds while the fast participant finished in only 68. This difference can be mainly attributed to the fact that the slow spent more time reflecting between transactions (2.5 times more than the fast participant).

The time between transactions consisted of two activities: reflection (looking around planning future moves) and what we call faulty transactions: failure to complete an attempted move of a slide from one position to another. The faulty transactions occurred when the participant dropped a slide while dragging it to its destination (2 instances), the participant selected more than one slide (3 instances), and the participant selected or moved an unintended slide (5 instances). Often the participant did not retain steady pressure, and so unwittingly trans-

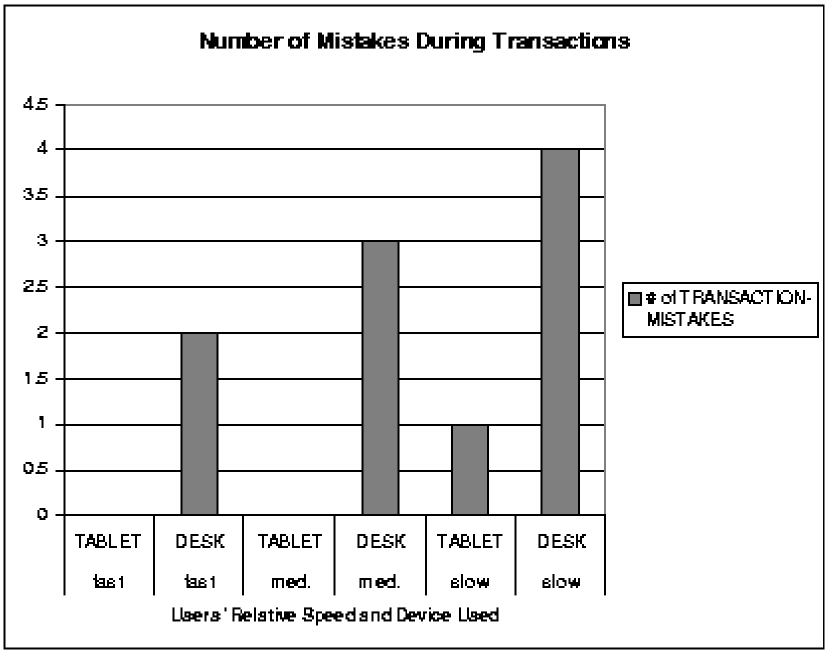

Figure 6
Number of faulty sorting transactions made on the desk and tablet by three representative participants.

ferred the selection from the intended slide to one that happened to lie in the trajectory through which the slide was being moved. This would result in the original slide "jumping back" to its original starting point, thus requiring another attempt to move it, a difficulty associated with dragging in early studies of light pens (Greenstein and Arnaut, 1988). None of these participants left any shapes in the wrong position when they finished the task, so these process faults are unrelated to the sorting errors described in Figure 3. Among these three participants, far more faults were made on the desk than on the tablet, as shown in Figure 6.

By definition faults increase the selection time, but they also seem to have a lingering effect on the participant. On average, participants needed more recovery time after making a fault on the desk. Even after successfully recovering from the fault, the following selection period was still an average of 159 milliseconds above the mean selection period of 2.3 seconds.

3.3 Fatigue and Learning

Despite the mental strain of trying to complete a difficult task quickly, and the physical strain of making large arm motions to sort on the desk, there was no evidence of fatigue between the first and last environment. This is surprising

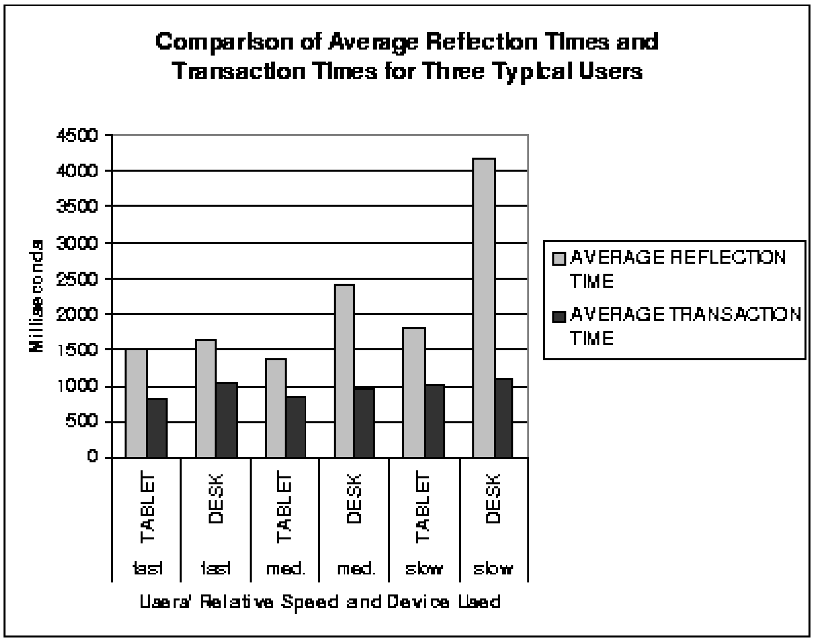

Figure 7
Comparison of average reflection times and transaction times of three representative participants.

because fatigue is frequently cited as an undesirable side effect of pointing over a display (Shneiderman, 1998). Fatigue is associated with applying pressure to maintain selections (Greenstein and Arnaut, 1988), an observation relevant to our research because the tablet requires more pressure than the desk to activate the stylus, while the desk requires less pressure over greater distances and longer periods of time. If fatigue were a factor, we would expect the time needed to complete the sorting task to increase as participants progressed from the first to the third device. It is also possible that although the participants sorted a unique set of images every time, they might have developed strategies for sorting more efficiently. In this case, we would expect the times to decrease between the first and third trials. Our data supported neither proposition; there was no statistically significant difference between the first and third tasks ($p = 0.44$). Because people less than 5'6" are potentially more vulnerable to fatigue due to the extra effort required to reach the periphery of the desk, we examined if this group separately. For people less than 5'6" there is weak evidence of fatigue ($p = 0.33$), but still without statistical significance. This result may be due to the relatively short duration of the tasks.

As noted above, the monitor with mouse was generally preferred for the image sorting task. We suspect the reason for this is twofold. First, the mouse is more accurate than either of the pens at successfully selecting the slides and transferring them from one position to another. The pen requires the participant to maintain continuous pressure while simultaneously dragging it across the screen. The mouse was also more accurate at selecting a slide in question; often a pen tap failed to select the desired slide.

MacKenzie, Sellen, and Buxton (1991) report that the mouse is superior for dragging and the tablet and pen is superior for pointing, reporting on a comparative study of pointing and dragging done with a mouse, stylus, and trackball.[6] This effect may be heightened in our study by PowerPoint's apparent optimization for mouse-style clicking and dragging.

The monitor may also have been preferred because its screen size may be more appropriate for the sorting task than that of either the desk or the tablet. The tablet had the disadvantage of showing very small images, while the desk had the disadvantage of requiring the movement of the images across long distances. The intermediate size of the monitor may be a more appropriate tradeoff between image visibility and distance necessary to be moved. to this observation is that it assumes the need for frequent access to the periphery of the display.

3.4 Sorting Task: Qualitative Results

Despite our initial hypothesis that the desk would be the architects' environment of choice for image sorting, only 3 of the 21 participants preferred it. A review of qualitative differences between the desk and other devices produced two potential sources of dissatisfaction: slow speed and poor resolution.

Slow Speed. Once we learned that shorter height was correlated with slower time and dispreference for the desk, we explored if the inverse was also true: were the participants who preferred the desk faster at sorting on it? We found no difference in the total desk sorting times between those participants who preferred the desk and those who did not ($p = 0.48$, ANOVA). Although shorter people were both slower and dispreferred the desk, those who did prefer the desk were no faster on it. Comparing the sorting times of those who preferred the tablet with those who did not yielded no statistically significant difference. The finding is the same for the monitor. From this we conclude that speed does not predict preference, although slow speed correlates with dispreferring the desk. The people who preferred the desk were tall but slow; speed did not effect their satisfaction with the desk.

Screen Resolution. A second limitation of the desk that could contribute to dissatisfaction is the relatively poor resolution. An intuitive understanding of how designers work with photographs would suggest that higher resolution is important, and that poor resolution would be a limiting factor in viewing photographs.

The desk has only 1626 pixels per square inch, compared with 6425 pixels

6 This study measured use of a pen on a special pad similar to a mouse pad, not a pen used directly on a display surface. The small, localized movements of a pen or mouse on a pad are quite different than the long, continuous movement of a pen across the surface of the display.

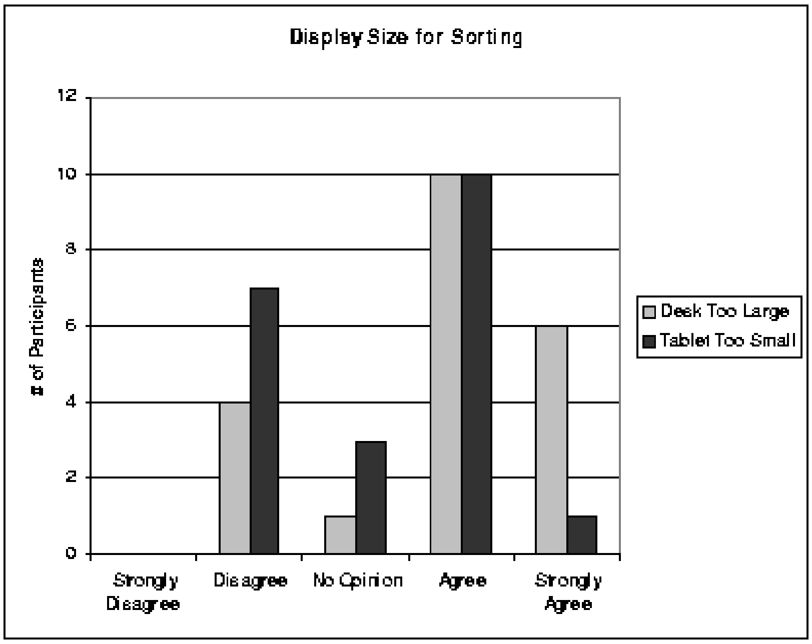

Figure 8
Participants' comments on the display size of the digital desk and tablet.

per square inch on the monitor, and 8411 pixels per square inch on the tablet. Because the desk has a lower maximum resolution than the other two devices, we had to chose between forcing the other devices to display at a lower resolution than they were capable of, or setting each device to its maximum resolution and allocating the same proportion of the available display to each image. We chose the latter approach, so the larger display resulted in physically larger images. Images were 4" x 2.9" on the desk (79 x 66 cm), 2" x 1.4" on the monitor (43 x 30 cm), and 1.43" x 1.06" (28 x 21.6 cm) on the tablet.

Interestingly, the participants did not report that the relatively poorer resolution on the desk to be important. Even though the desk has 5 times fewer pixels per square inch than the tablet, fewer participants said the desk's display was "hurt their eyes" or was "hard to read" than they did for either the tablet or the monitor. This suggests that a relatively larger image size might make up for a relatively poorer resolution in terms of user satisfaction.

3.5 Display Size

Most of the participants in this study felt the desk was too large for sorting, and half felt the tablet was too small. The monitor, the middle-sized device, is the preferred device for image sorting. As Figure 8 shows, the consensus about the desk being too large and the tablet too small suggests that an intermediate size is the most appropriate. However, preference for the monitor's size may have been influenced by the use of the mouse, since the distance the hand must move when using the mouse with the 19" monitor is usually less than what is needed for the corresponding move using a pen input device. Nevertheless, since over half of the respondents said the tablet was too small, we can assume that relative image size does play a role.

An alternative to shrinking the size of the desk's display to be closer in size to the monitor would be to separate the usage of the space into a primary working area and a periphery which holds contextual information, in much the way that physical desks are used both as a work surface and as a storage area for piles of documents that are accessed only occasionally (Mackinlay, Robertson, and Card, 1991). The study required the participants to access the entire surface actively, reflecting the fact that most commercial software uses the entire display. This point is revisited in the conclusions section below.

3.6 Support for Sorting on the Desk

Although only three participants preferred the desk for image sorting, they made up in enthusiasm what they lacked in numbers. All three had strong feelings that the desk was superior. Their enthusiasm for sorting on the desk is marked because the majority of the participants, 15 in all, either agreed or strongly agreed that the desk was too large for sorting. (Three participants did not express strong feelings in either direction.) This reflects a trade-off between the benefits of relatively large images and the liabilities associated with the large space required for operating on the images.

Given the problems of slower speed, poorer resolution, and less easily reached images, why would anyone prefer the desk for sorting? Characteristics of the three participants who did prefer the desk include: height greater than 5'6", little concern for speed, and a manner of expressing enthusiasm for the desk by making wide sweeping motions with their arms. To explore what might account for this preference for the desk we reviewed these participants' comments during the open-ended interview questions. The following are responses to "Why is the desk your favorite environment for sorting?"

> "Standing at the desk made it easier to take control of the environment." "Working on the desk is better ... it's more like picking things up." "Freedom from sitting is great. I can think ... I'm an architecture person and I need that freedom to work."

From these comments we can gather that for these few participants, the reason for preferring the desk stems from a similarity to the traditional environment for architectural work: the drafting table. Although the study results do not support our initial hypotheses that in general the participants would prefer the desk because it is more like the traditional environment for architectural information retrieval tasks, the few who did prefer the desk apparently did so for that reason.

3.7 Sketching Task: Preferences

Although relatively few participants preferred the desk for image sorting, half of them did prefer it for sketching, with 10 preferring the desk, 10 the tablet, and 1, a former engineer with extensive CAD (Computer Aided Design) experience, preferring the monitor and mouse. Because of the near-even split in device preference for sketching, we wanted to understand what characteristics were shared by those who preferred the desk versus those who preferred the tablet.

Sketching has an emotional association for architects; this activity lies at the heart of how they are taught to design. Sketching is also a "practice" in that like music, it is necessary to do a little every day to stay proficient. The highly personal nature of sketching, formed by training and daily personal habits, lends itself to qualitative analysis of preference. Therefore, we organized our inquiry into device preference for sketching around an expectation that preference for a computerized sketching environment would correlate with preferences for a non-computerized traditional sketching environment.

We evaluated sketching on the desk from two perspectives: user satisfaction with the sketching experience, and outside experts' opinions on which sketches were better. We gathered background information about the participants' architectural training and weekly work habits through a written survey. We expected to see a correlation between years of architecture experience and preference for the desk. The data did not support this claim and years of architecture experience or even hours per week spent sketching had no relationship to device preference, ($p = 0.4$, ANOVA). We also expected that participants who spent more hours per week working standing at a drafting table would prefer to work standing at the desk. The data did not support this claim either ($p = 0.36$); there were no correlations with any of the seating options and preferring the desk for sketching.

After participants completed the sketching task, we asked them to fill out a written survey asking which factors contributed to their preference for a particular device. The survey asked about their preferences for display size and input device for sketching. When asked if input device or display size is more important, 19 of the 21 participants said that input device is more important than display size for sketching. However, this preference did not correlate to a preference for a particular display. Because the participants did not differentiate input devices from displays, we relied on responses to the open-ended interview questions to learn about how the affordances of display size affected the sketching task.

3.8 Display Size

Of the 10 who preferred sketching on the desk, 4 specifically mentioned liking the larger area with comments such as:

> "Sketching has a larger gesture, it's hard to draw on the tablet... too confining." "It's more comfortable (with) more space... it's like how I learned to produce architecture."

But the smaller size of the tablet brought favorable comments from 5 of the 10 who preferred the tablet. Two of the participants said explicitly that the desk

	Pen with Digital Desk	Pen with Tablet
Weight	Light	Heavy
Thickness of “ink”	Pencil width	Magic-marker width
Pressure required to make a mark	Light	Heavy
Stylus tip – cursor alignment	Poor	Good

Table 3
Comparison of the relative qualities of the pens for desk and tablet.

is too large for sketching because it makes it difficult to capture the proportions of a drawing. Three others mentioned that the tablet felt the most like their 8.5” x11” sketch pads that they sketched on in the field, or that they usually sketched on scraps of paper. The tablet’s lighter weight appealed to them, but in general the participants handled it more gingerly than necessary. Comments about how fragile the tablet seemed were common, and many felt they were in danger of breaking the screen by pressing down hard while drawing. A few tried to put the tablet in their laps to draw, but ultimately put it back on the desk, commenting that it didn’t feel stable resting on their laps.

3.9 Pen Alignment

These comments about the advantages of large and small display sizes are far less cohesive than the relatively consistent comments about the input devices. The tablet and the desk required different pens for input, even though all of the hardware was made by ITI-VisionMaker, the same manufacturer. The differences between the two pens are summarized in Table 3.

The participants responded differently to the desk and pen versus the tablet and pen. In addition to agreeing that input device is more important for sketching, there was also general agreement that the pen used with desk had an alignment problem. Poor pen alignment on the desk was commonly mentioned as a reason for preferring the tablet. Additionally, 3 participants who preferred the desk for sketching said during the interview that what they really wanted was to use the pen from the tablet on the desk.

Pen misalignment is a perceived difference between the location of the stylus’ tip and the location of the cursor, caused by parallax errors. The digital desk uses a projected signal, and the thickness of the screen surface is perceptible to the user. The problems of parallax errors are described by Greenstein and Arnaut (1988) in a study of touch screen displays. They report user frustration with target selection when the touch points are not centered over the targets. They also identify parallax errors as a problem with light pens, writing:

> “Parallax may be a problem, especially when pointing to objects at the sides of the display. Using only the center of the screen to display targets may alleviate this problem; however, this limits the usefulness of the device. A more satisfying solution might be to make targets at the sides of the screen large enough to minimize the effects of incorrect placement.” (Greenstein and Arnaut, 1988), p. 501.

Parallax problems, or the apparent displacement of an object as seen from various points of view, are common to all projections from 2D (the cursor) into 3D (the display surface). In response to this, Ullmer and Ishii (1997) whose metaDESK work uses the same digital desk device, mitigate the problem by creating a 1 cm thick "passive lens" rimmed with wood to give the illusion of lifting the display surface, thus masking the alignment errors. This approach works provided the display surface is horizontal, but because of our desire to mimic the work context traditional to architecture, we did not want to constrain the angle of the display surface. (And the desk surface used in this study could not be made to lie less than 12 degrees above horizontal.)

VisionMaker provides alignment utilities for the desk and tablet that the participants could use to register the tip of the pen. However, the alignment utility could not accommodate the variability in participants' styles of holding the pen. Many participants thought the pen alignment was fine after using the alignment utility, but then became frustrated while sketching if they tended to vary the angle of the pen. The pen alignment varies equally with the pen angle for all users, but some participants tended to vary the pen angle more than others. This is partially a matter of drawing style and partially of instruction. Architects who learned to draft manually hold an ink pen vertically to make a straight line without smudges, while some designers make rough sketches using fine art methods of holding a soft pencil nearly horizontal, sometimes even with the index finger on top of the pencil.

When sketching, participants were not permitted to erase any lines, but they could start over as many times they wished. In practice no one used more than two trials, but negative comments about the alignment such as "This is too messy" were common. Frustration with the alignment was common among all participants, even those who preferred the desk. Typical comments about the pen with the desk include:

> "Pen tracking is critical, and that pen (with the desk) failed."
> "The funny angle is awkward... why does it do this?"

We believe the different responses to pen alignment on the desk are due to variation in individual drawing style, with the participants using the greatest range of pen angles becoming the most frustrated. Three of the four participants who clearly stated they preferred the tablet to the desk because of superior pen alignment used the greatest range of pen angles in the study. The fourth user with a wide variety of pen angles also preferred the tablet for sketching, but did not mention poor pen alignment as a motivating factor for the choice.

3.10 Sketching Task: Judges' Assessments

The participants' frustration with the desk's poor pen alignment was reflected in the judgements of the five professional designers who reviewed the sketches. All 63 sketches made during the study were printed out on 8.5" x 11" paper to disguise on which device the sketch was created. The Web Gallery feature of

Sketch Description		Votes by Judge #					
Rank	Created Using	J-1	J-2	J-3	J-4	J-5	Total
1 (tie)	Tablet	1	2	2	1	2	5
1 (tie)	Tablet	3	1	2	0	2	8
3	Tablet	1	0	2	1	3	7
4	Tablet	1	1	1	3	0	6
5	Digital Desk	0	1	1	1	1	4

Table 4
Rank of sketches that received the most votes along with distribution of judges' votes.

Adobe Photoshop was used to standardize the presentation. This feature creates a standardize image size with the content centered.

Evaluating the sketches required two passes. Initially the judges were told they were looking at a stack of drawings made on a computer, and were asked to fill out an evaluation form selecting their favorites and explaining why. They kept the packets for two weeks and left to work independently. This method of evaluation did not yield any usable feedback. The judges rebelled by not filling out the evaluation forms and saying they couldn't evaluate technical expertise over artistic merit. After this, the experimenter sat down with the judges one-on-one for an evaluation period. Each judge was given a set of stickers corresponding to medals for "best of show." Each judge was asked to create categories, which were conglomerated by the experimenter into four categories: (i) best likeness to original, (ii) confidence, (iii) expressiveness, (iv) scale. For each category, the judge assigned three unranked best of show awards.

The distinctions between the categories were not rigid, but rather were used as inroads to encourage the judges to expose the rationale for their evaluations. Thus they should not be taken as a generalizable set of criteria for evaluating the effectiveness of pen based interfaces. To this group of judges, "best likeness to original" meant the closest thing to a photocopy of the original sketch; "confidence" meant the lines were straight and firm with neat corners, not shaky or hesitant; and "expressiveness" meant that the sketch didn't seem like a copy, and the judges felt they were looking at something the architect wanted to be making rather than a copy. "Scale" meant maintaining the same size figures and space between figures as the original. For example, a copy made on a photocopier set to 90% reduction would produce a smaller result, but the proportions would be the same. Similarly, a sketch could be voted for accurate scale, even if the figures in the sketch were messy; the important aspect is that the relationship between the figures is not warped or distorted.

Although each judge created categories independently while talking with experimenter, the judges exhibited a fairly high degree of consensus. The total votes are summarized in Table 4. Twenty-one of the 63 drawings received at least one vote. Six participants had more than one sketch receive votes, but only 4 participants received 8 or more votes. These 4 participants each created one of the top 4 sketches, receiving 6 or more votes each across categories.

The four most-awarded sketches have more in common than being created by the most artistic participants: they were all made using the tablet. This indicates

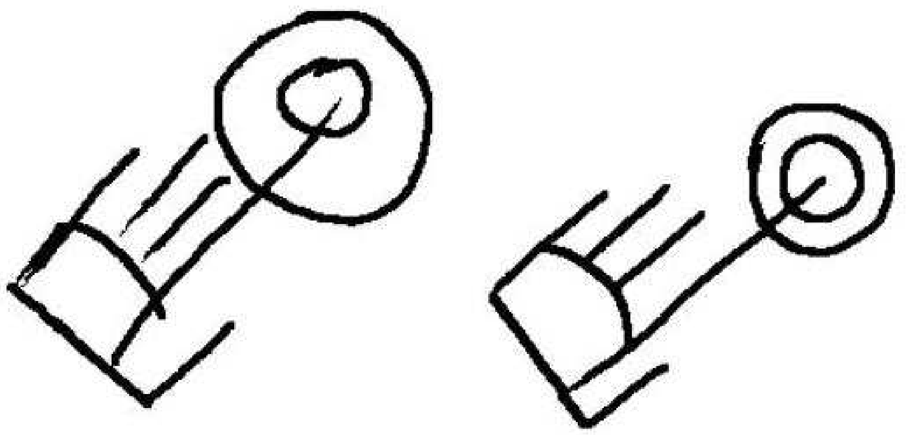

Figure 9
Sloppy intersections in a figure created on the desk (left) compared with one created on the tablet (right).

that the sketches, not the participants, were being evaluated. In the final tally of votes for sketches by device, the tablet received 68% of the total votes, the desk 30%, and the monitor, with 1 vote, 2%. Despite the judges' overwhelming preference for sketches made on the tablet, a casual observer would not be able to determine which sketches had been made on the tablet and which on the desk. However, sketches made on the monitor do stand out because of their blocky curves and crossing and re-crossing of the same line.

Two potential reasons why the tablet might be more successful than the desk are its superior pen alignment and its display size which is similar to that of a standard-sized sheet of paper. Although in the initial evaluation the judges had a difficult time talking about why some sketches were better than others, many absent-mindedly made marks on the sketches, circling places with sloppy line intersection. Evidence of pen misalignment shows up in the sketches when a participant is trying to close a figure, such as a square. The architect will lift the pen from the table at the moment she sees the stylus tip touching, but not crossing, the line she wants to meet. Because parallax errors cause the cursor to misalign with the tip of the stylus, the likely result is that the line extends to over-cross its destination, or does not extend far enough and fails to meet the target, as seen in Figure 9.

Sketches exhibiting over-crossing were repeatedly penalized in the categories of confidence, expression, and likeness to original, all of which required crisply defined line intersections in the sketch. The only category that didn't take pen alignment into account is scale, but as mentioned above, the close match between the tablet's display size and the size of a standard sheet of paper gave the tablet an advantage in the scale category as well.

The scale category favors the tablet because the drawing surface is the same size (11" x 8.5") as the original sketch the architects copied. The large display surface on the desk makes it more difficult to calculate how big to make the figures. For example, in an arrangement that requires three 2" tall figures to fit vertically, the participant drawing this particular group of figures on the tablet has only to locate them within the 8.5" height of the tablet, easily done by "eyeballing" the display surface into thirds. On the tablet, making one figure 1/2" larger than it should be will use up 20% of the remaining space on the drawing surface and risk crowding the other figures. However, someone making the same mistake on the desk will use only 2% of the remaining space on the desk, a difference much more difficult to detect.

The judges felt that that the scale category demonstrated a different kind of skill than the others, for two reasons. First, this category allowed drawings with poor line quality the opportunity to get votes; and second, maintaining the scale of a copied drawing is an important architectural task. When architects make visual notes in their own sketchbooks from an image or from a real-life scene, the important piece of information is a description of the relationship between the elements in the image; the relationship is generalizable and applicable to other situations, the specific qualities of the forms depicted are less important.

The scale category is closely related to the "best overall likeness" category because proportion is one aspect of a good likeness. Although we can report that display size and pen alignment contribute to the reviewers' categories, the overlap between the categories underscores the difficulty of making judgments about something as subjective as sketches. Subjective preference is critical to the design of image manipulation systems because although we can conclude that pen alignment is an important part of sketching both for participants and for outside reviewers, overall preference for sketching device overall is more complex. Despite all winning sketches being created on the tablet, half of their creators preferred sketching on the desk rather than the tablet, the same proportion as the study as a whole.

3.11 Limitations of the Study

This study has some limitations that may have affected the outcome and dampened architects' enthusiasm for the desk. As discussed above, the radically different alignment performance of the pen for the desk may have undermined our attempts to isolate the effects of display on these tasks. The users had difficulty distinguishing between preferences for a display size and preference for the input device that went with that display. Additionally, the choice of PowerPoint may have impacted performance on the image sorting task because we suspect that

PowerPoint is optimized for use with a mouse. Our study was limited because in an attempt to isolate the effect of screen size, we had participants sort images in a rigid tiled arrangement, instead of one allowing for overlapping piles as seen in practice. A final limitation is the weaknesses of the data observation. For example, because the participants shifted position as they worked, the camera was not always positioned properly to capture all of their actions. Additionally, only one camera was used, recording the participants' action on the screen, but not showing head movement or gaze direction. Individual transaction times were recorded from the time stamp of the videotape, increasing the margin of error in data collection.

4. Related Results

Ergonomic concerns related to large displays and body dimension have been addressed in other research projects involving large displays, but primarily in terms of interaction techniques. A common problem addressed in these projects is the difficulty of selecting commands from menus located far away from the user (Mynatt et al., 1999; Streitz et al., 1999). The most common solutions are to use gestural input (Long, Jr., Landay, and Rowe, 1999) or marking menus (Kurtenbach and Buxton, 1991), which have the advantage of allowing the user to access command operators from any part of the screen. We, however, are interested in object selection and movement across a large display.

Kurtenbach and Buxton (1991), in their GEdit system, explain users' interaction with the pen by moving a group of objects along a trajectory as a "single short continuous movement." In this system, users depress the stylus, then execute the command, and then release the pressure. This approach, though valuable for gestural commands, seems unlikely to scale to long continuous movements as required by large surfaces. Early critiques of light-pen interfaces point out a similar problem (Greenstein and Arnaut, 1988), suggesting that light pens be implemented with two modes: pointing and tracking. Having two modes avoids the difficulty of maintaining a steady rate of cursor movement, which is necessary to avoid interrupting tracking and losing the cursor. Flatland and Tivoli use standard selecting and dragging with the pen for allowing users to, for example, re-order items in a list. They do this in the interest of maintaining naturalistic interaction, but for the reasons described above, we feel this might not be the best use of pen-style interaction on a large display surface.

An alternative interaction technique implemented by Rekimoto and Saitoh (1999) is the pick-and-drop technique. This is specifically designed to avoid the problem of keeping a stylus tip depressed while dragging over a large area, although it is intended for multi-device interaction. In this case a user uses a pen to select a file to transfer by tapping and then releasing pressure from the screen. Pen ID's are used to hold the data and when the pen moves over the surface of another display a virtual shadow of the held object gives the user feedback about where to drop on the target. This kind of interaction seems like a better alternative for a single device with a large display, but more study is needed because it has not been systematically applied to repetitive tasks like image sorting.

Another ergonomic concern is the amount of head-turning required to take in the entire surface of a large display. Emerging interest in eye-tracking and head-mounted displays could provide insight into how users scan the displays to plan their next interaction with the system (Jacob, 1995).

5. Conclusions

The results of this study do not support our initial hypothesis that architects would prefer completing image design tasks on the Digital Desk; the desk was strongly dispreferred for image sorting and only partially preferred for sketching. The desk did only as well as the tablet in qualitative evaluation of the sketching tasks and much worse than either the tablet or the monitor in image sorting. Quantitative results show more sorting mistakes on the desk than other environment and significantly longer times to complete the image sorting task on the desk.

This study provides evidence that pen performance is critical for architects' satisfaction with a computer-based sketching environment, and the poor pen alignment on the desk is a significant obstacle to its acceptance by architects. The conclusions that pen alignment is critical for artistic drawings seems self-evident, but this study is unique in that it presents results from a user study, rather than anecdotes from the development cycle. Commenting on the cumbersome nature of pen-based interaction is also not new, but this article describes issues that arise in an empirical study with real potential users.

Further research is needed to develop more appropriate interaction techniques for image manipulation on a large surface. It is unclear if a pen with improved interaction techniques would be preferable to a mouse for an intense manipulation task like image sorting, but clearly a transfer of mouse-style dragging interaction to pens for large displays is not desirable. Rather than having to maintain steady pressure to move objects, pen-based direct manipulation interfaces should allow users to select an object by tapping on it and deposit it in a new location by tapping again, somewhat like a simpler version of cut and paste operations. These techniques could expand on the pick-and-drop technique (Rekimoto and Saitoh, 1999). Although gesture-based commands could be helpful for image browsing or sketching, they generally are not used for indicating both a beginning and an ending location. A touch-sensitive interface in which the hands can be used directly might also be preferable.

Although the desk's display produces a lower image resolution, the participants did not indicate that this was problematic. The primary liability of the projected display was the parallax problem resulting from an increased distance from the displayed image to the interaction surface. This has interesting implications for technology development because it suggests that slimming the display surface to improve pen-alignment should take precedence over increasing the resolution.

Previous authors have suggested that it is important to make a distinction between the information that is the current focus of attention and the information that is only peripheral to the current tasks (Furnas, 1986; Mackinlay, Robertson, and Card, 1991). However, these arguments have been made primarily from the cognitive perspective – that this kind of organization helps users better understand and organize the information they are using. The unexpected finding in this study that height influences preference and sorting speed for the digital desk supports this notion for physical reasons as well. Since the original motivation for the focus/periphery distinction was physical placement of artifacts on a desktop, this result completes a cycle of reasoning.

Most of the participants in this study felt the desk was too large for sorting, and half felt the tablet was too small. The monitor, the middle-sized device, is the preferred device for image sorting. The consensus about the desk being too large and the tablet too small suggests that an intermediate size is the most appropriate. However, preference for the monitor's size may have been influenced by the use of the mouse, since the distance the hand must move when using the mouse with the 19" monitor is usually less than what is needed for the corresponding move using a pen input device. Nevertheless, since over half of the respondents said the tablet was too small, we can assume that relative image size does play a role.

It is worthwhile considering if the desk might be made more suitable for architectural design tasks if some adjustments were made, particularly improving the pen alignment, changing the manner in which the pen is used for moving objects long distances, and making a smaller version of the display available for smaller users. It might also be useful to explore how displays of different sizes could work together as a system. Just as architects use different sizes of paper for different tasks, they may want varying sized computer displays. The tablet could be used for mobile tasks and then information could be transferred to the desk. The large surface of the desk could be better used to support architectural practice by having reminder images in the periphery help architects link current sketches with previously retrieved images.

Although many aspects of how to build computer environments that support architects' image manipulation tasks are not well-understood (Elliott, 2000), supporting the practices that have worked effectively for architects in the physical world seems a good way to begin. Although responses to the digital desk were not as positive as we believed they would be, some of the participants in the study did specifically mention that it provided a better fit with their customary work habits. We believe that with flexible display size, use of the periphery for peripheral information, and improved pen-based interaction techniques for large surfaces, digital desks could become effective tools for supporting architects' image manipulation tasks.

Acknowledgments
We thank Jonathan Henke for his help in conducting usability studies. This research was supported in part by NSF Research Infrastructure Award EIA-9802069.

References

Accot, Johnny and Shumin Zhai. 1999. Performance evaluation of input devices in trajectory-based tasks: An application of the steering law. In Marian G. Williams, Mark W. Altom, Kate Ehrlich, and William Newman, editors, *Proceedings of the Conference on Human Factors in Computing Systems (CHI-99)*, pages 466–472, New York, May 15–20. ACM Press.

Blau, J. 1984. *Architects and Firms*. MIT Press, Cambridge, MA.

Cuff, D. 1991. *Architecture: The Story of Practice*. MIT Press, Cambridge, MA.

Elliott, Ame. 2000. Computers for architects. Technical Report CEDR, Working Paper Series, UC Berkeley.

Fitzmaurice, George and William Buxton. 1997. An empirical evaluation of graspable user interfaces: Towards specialized, space-multiplexed input. In *Proceedings of ACM CHI 97 Conference on Human Factors in Computing Systems*, volume 1 of *PAPERS: Handy User Interfaces*, pages 43–50.

Furnas, George W. 1986. Generalized fisheye views. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 16–23. ACM, April.

Greenstein, Joel S. and Lynn Y. Arnaut. 1988. Input devices. In Martin Helander, editor, *Handbook of Human-Computer Interaction*, number 22 in II. User Interface Design. North-Holland, New York, NY, pages 495–519.

Guimbretiere, Francios, Maureen Stone, and Terry Winograd. 2001. Fluid interaction with high-resolution wall-size displays. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2001)*, Orlando, FL, November. ACM.

Jacob, R. J. K. 1995. Eye tracking in advanced interface design. In W. Barfield and T. A. Furness, editors, *Virtual Environments and Advanced Interface Design*. Oxford University Press, New York, NY, pages 258–288.

Kalay, Y. E. and G. Carrara. 1994. *Knowledge-Based Computer-Aided Architectural Design*. Elsevier Science Publishers, Amsterdam, The Netherlands.

Koutamanis, A., H. Timmermans, and I. Vermeulen. 1995. *Visual Databases in Architecture: Recent Advances in Design and Decision Making*. Avebury Press, Aldershot, UK.

Kroemer, K. H. E. 1988. VDT workstation design. In Martin Helander, editor, *Handbook of Human-Computer Interaction*, number 23 in II. User Interface Design. North-Holland, New York, NY, pages 521–539.

Kurtenbach, Gordon and William Buxton. 1991. Issues in combining marking and direct manipulation techniques. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Input Techniques, pages 137–144.

Larson, M. S., G. Leon, and J. Bolick. 1983. The professional supply of design; a descriptive study. In J. Blau, M.E. La Gory, and M. S. Pipkin, editors, *Professionals and Urban Form*. State University of New York Press, Albany, NY.

Long, Jr., Allan Christian, James A. Landay, and Lawrence A. Rowe. 1999. Implications for a gesture design tool. In Marian G. Williams, Mark W. Altom, Kate Ehrlich, and William Newman, editors, *Proceedings of the Conference on Human Factors in Computing Systems (CHI-99)*, pages 40–47, New York, May 15–20. ACM Press.

MacKenzie, I. Scott, Abigail Sellen, and William Buxton. 1991. A comparison of input devices in elemental pointing and dragging tasks. In *Proceedings of ACM CHI'91 Conference on Human Factors in Computing Systems and Graphics Interface*, pages 161–166. ACM-Press.

Mackinlay, Jock, George Robertson, and Stuart K. Card. 1991. The perspective wall: Detail and context smoothly integrated. In *Proceedings of the ACM CHI '91 Human Factors in Computing Systems*, pages 173–179.

Mynatt, Elizabeth D., W. Keith Edwards, Anthony LaMarca, and Takeo Igarashi. 1999. Flatland: New dimensions in office whiteboards. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems*, volume 1 of *Profiles, Notes, and Surfaces*, pages 346–353.

Newman, William and Pierre Wellner. 1992. A desk supporting computer-based interaction with paper documents. In *Proceedings of ACM CHI'92 Conference on Human Factors in Computing Systems*, Desks, Video, and Screens, pages 587–592.

Pedersen, Elin Ronby, Kim McCall, Thomas P. Moran, and Frank G. Halasz. 1993. Tivoli: An electronic whiteboard for informal workgroup meetings. In *Proceedings of ACM INTERCHI'93 Conference on Human Factors in Computing Systems*, Meetings and Collaborative Writing, pages 391–398.

Rekimoto, Jun and Nobuyuki Matsushita. 1997. Perceptual surfaces: Towards a human and object sensitive interactive display. In *Proceedings the of Workshop on Perceptural User Interfaces (PUI'97)*.

Rekimoto, Jun and Masanori Saitoh. 1999. Augmented surfaces: A spatially continuous work space for hybrid computing environments. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems*, volume 1 of *Augmented Surfaces*, pages 378–385.

Shneiderman, Ben. 1998. *Designing the user interface: strategies for effective human-computer interaction*. Addison-Wesley.

Stevens, G. 1998. *The Favored Circle: The Social Foundations of Architectural Distinction*. MIT Press, Cambridge, MA.

Streitz, Norbert A., Jorg Geissler, Torsten Holmer, Shin'ichi Konomi, Christian Muller-Tomfelde, Wolfgang Reischl, Petra Rexroth, Peter Seitz, and Ralf Steinmetz. 1999. i-LAND: An interactive landscape for creativity and innovation. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems*, volume 1 of *Collaborative and Multimedia Systems*, pages 120–127.

Ullmer, Brygg and Hiroshi Ishii. 1997. The metaDESK: Models and prototypes for tangible user interfaces. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Blurring Physical and Virtual, pages 223–232.