Lightweight Task/Application Performance using Single versus Multiple Monitors: A Comparative Study

Youn-ah Kang and John Stasko

School of Interactive Computing / GVU center
Georgia Institute of Technology
85 5th Street, NW, Atlanta, GA 30332
ykang3@mail.gatech.edu, stasko@cc.gatech.edu

ABSTRACT

It is becoming increasingly common to see computers with two or even three monitors being used today. People seem to like having more display space available, and intuition tells us that the added space should be beneficial to work. Little research has been done to examine the effects and potential utility of multiple monitors for work on everyday tasks with common applications, however. We compared how people completed a trip planning task that involved different applications and included interjected interruptions when they worked on a computer with one monitor as compared to a computer with two monitors. Results showed that participants who used the computer with two monitors performed the task set faster and with less workload, and they also expressed a subjective preference for the multiple monitor computer.

CR Categories: H.5.2 [Information Interfaces and presentation (e.g., HCI)]: User Interfaces - Evaluation/methodology

Keywords: Multiple monitors, large display, evaluation, window management

1   INTRODUCTION

The use of multiple monitors (displays) on computers is becoming more and more common. It is not unusual to walk through an office of information technology workers and observe many computers with multiple displays on people's desks. We speculate that a combination of decreasing monitor prices, the emergence of flat screen LCDs requiring less desktop area, improved operating system support, and cheaper, more powerful graphics cards has helped to fuel this increasing presence. Of course, these factors are likely combined with a fundamental desire in people for more screen real estate to distribute all the different windows and applications they use.

Correspondingly, research on the use and potential benefits of multiple monitors [9, 20] has been increasing as well, but we feel that it has not kept pace with multiple monitors' growing use. In fact, the majority of research in this area arguably has been on the development of new window management operations to benefit interaction and space management [2, 3, 4, 12].

Our intuition generally leads us to believe that multiple monitors should benefit productivity and better satisfy users. Having more space to work with is broadly viewed as more beneficial – people seem to like large (physical) work surfaces on which to spread out their materials. However, on computers, the increased screen area may force people to perform more window operations and perform more space management, so our intuition

actually may not be correct.

Unfortunately, there are few, if any, careful studies of single and multiple monitor computer usage to compare task performance, cognitive load, and user satisfaction. Furthermore, research studies are needed that help explain why (or why not) multiple monitors are actually beneficial and if there is a benefit, what its magnitude is. This research also may identify ways that we can help people make even better use of the increased screen area that multiple monitors provide.

This article reports on a comparative study of people performing lightweight everyday tasks on a computer with either one or two monitors. Participants planned a business trip and interacted with familiar desktop applications such as word processors, web browsers, and spreadsheets. We studied completion times, cognitive load, and subjective impressions of the participants. We also examined whether prior multiple monitor usage affected performance, and we observed participants' patterns of use under the two different configurations.

Experimental results showed that participants performed the task set faster and with less workload while using the multiple monitor computer. Participants also expressed a subjective preference for multiple monitors compared to a single monitor. Other factors, such as prior experience with multiple monitors and window management style, had lesser effects that were more dependent on experimental ordering issues.

2   RELATED WORK

Previous work relevant to our research includes (1) studies of multiple monitor usage and window management (2) evaluation studies on potential benefits of multiple monitor and large displays and (3) virtual multiple monitor use in control room situations..

2.1   Multiple monitor usage and window management

Fieldwork by Grudin [9] described the benefit of multiple monitor system qualitatively. He documented the usage patterns of CAD/CAM programmers and designers running multiple monitors. The study showed that an additional monitor is often used to hold windows that help the user maintain awareness, to display information relevant to the current task, or to keep communication applications (e.g., email) visible. Despite the limitations observed in OS support, multiple monitor users clearly liked the extra screen real estate, and they were willing to adapt their windows and application layouts optimally for the number, size, orientation and resolution of their displays.

In a quantitative study, Hutchings et. al. [11] compared window management practices of single-monitor and multiple monitor users. Several findings from that work are consistent with those of Grudin's work and suggest that multiple monitor users tend to view several windows without interacting with them, some of which are related to a current task and others that are related to other tasks.

Other studies of window management practices, but not necessarily multiple monitor usage, help inform our work as well. Hutchings and Stasko [13] interviewed office workers who used a variety of window managers and display configurations in order to understand their high-level space management practices. They found that people specifically hide window contents for a variety of reasons, use "empty space," and employ space to help remind. Their findings emphasize the importance of window management issues in using multiple monitors. They also observed that multiple monitors are not treated as simply "additional space" and that people tend to manage space within monitors. In his study of people who use overlapping window managers, Gaylin [8] videotaped 9 users for 22 minutes each in their natural work environments to assess what window operation they used the most. He used the results to build benchmark tasks for evaluating overlapping window managers.

2.2 Evaluation on benefits of large display and multiple monitors

Several studies to evaluate potential performance benefits of large displays have been conducted, with far fewer examining multiple monitor usage.

Simmons [17] conducted a study comparing performance on monitors (17 inch to 21 inch), with slightly differing resolutions. The study found productivity benefits of higher resolutions; people performed standard high-level desktop applications tasks faster with the and slightly higher resolutions, as opposed to the smaller monitors. Another study [6] also reported performance benefits from using larger display sizes in that large screens enhanced text processing.

Dillon et. al. [7] assessed the effects of display size on readers' manipulation, comprehension and subjective impressions. The study results revealed that levels of reading comprehension are noticeably higher in the large window condition and that screen size affects how readers interact with a text. Another study suggests that the perception of words while skimming a newspaper headline on a computer screen is affected by screen resolution and monitor size [14]. However, these studies did not examine display sizes larger than a 21'' diagonal display or multiple monitors.

Patrick et al found that users performed significantly better at remembering maps when using a large projection display as compared to a standard desktop monitor [16].

In order to further explore the effects of display size and distance, Tan et al. constructed an experimental design in which a small and large display presented the same images at equivalent visual angle [18]. They found performance increases on spatial orientation tasks involving static 2D scenes when users worked on the large display. In their follow-up study [19], they examined the effects of physical display size on path integration in 3D virtual navigation tasks. They demonstrated that users perform 3D navigation tasks requiring path integration more efficiently on large displays than on smaller ones.

Czerwinski et. al [5] conducted research to evaluate the benefit of large display surfaces and demonstrated that large displays can help users reduce task completion times. The studies showed a (significant) 12 percent performance benefit; that is, study participants accomplished a mix of typical office productivity tasks 12 percent faster with a large display. In addition to productivity benefits, the studies showed that larger displays improve users' recognition memory and peripheral awareness. They also showed that users were more satisfied with large displays than small displays.

Ball and North [1] conducted a lab study to examine the potential benefit of high-resolution displays on performing a basic perceptual task and navigation. The results suggest that high-resolution displays can be a benefit in that they significantly improve performance time for basic visualization tasks in finely detailed data. They also found that the high-resolution displays help people find and compare targets faster, feel less frustration, and have more of a sense of confidence about their responses.

Tan and Czerwinski [20] reported a study examining the effects of visual separation and physical discontinuities when distributing information across multiple displays. Interestingly, they saw no effects of physical discontinuities on performance.

2.3 Tradeoffs between multiple monitors and multiple windows in other types of systems

While most studies on multiple monitors involve a normal work setting, another study investigated the possibility of multiple monitors in cramped environments such as cockpits and shipboard control rooms, with more emphasis on human factors and ergonomics [14]. The study examined virtual monitors using head-mounted display technology and compared the value of multiple monitors to that of a single monitor. Results showed that accessing information distributed across an array of monitors was at least as effective as locating and managing windows on a single monitor and that a task that requires only infrequent monitoring can be moved to a secondary, peripheral monitor without disrupting performance on current tasks.

3 COMPARATIVE STUDY

The primary goal of this study was to examine the effect of multiple monitors (on a computer) on people's performance of a set of common computing tasks. Also, we wanted to learn if and how people worked on the tasks differently on a computer with multiple monitors and whether prior experience with such display configurations, or the lack thereof, leads to different use and performance. We hypothesized that multiple monitors would lead to faster task completion and reduced cognitive workload, and that prior experience with multiple monitors would be beneficial.

3.1 Participants

We recruited 28 participants (7 females). All were either students or university employees and the average age of a participant was 26.4. All had at least moderate competency with Microsoft (MS) Windows and Office applications that would be used in the study, and 22 participants had experience with online travel agencies such as Travelocity and Expedia that also would be used. 12 participants were regular users of a multiple monitor computer.

We report findings from 27 participants; one participant did not use the second monitor at all under the multiple monitor condition and we removed that person's data from the analysis. The experiment took approximately one hour and participants were given $10 for their time.

3.2 Design

We conducted a within-subjects experiment in which each participant completed two sessions, one with a single monitor computer (Singlemon) and one with the two-monitor computer (Multimon). We will call this within-subject variable "Setting." In the Singlemon setting, participants used a computer running MS Windows XP with one 17" LCD monitor at 1024x768 resolution. In the Multimon setting, two monitors identical to the one used in the Singlemon setting were used. They were positioned side-by-side at an angle of 160° as shown in Figure 1.

Participants were randomly assigned to one of two experimental groups. 13 participants used the single monitor computer first and 15 used the two-monitor computer first. Each group used the alternate configuration in their second session.

Figure 1. Multimon configuration used in experiment: Two 17'' LCD monitors angled at 160˚, each at 1024 x 768 resolution.

We designed two sets of isomorphic tasks so that we could minimize the variance across the sessions. One set involved a trip to Boston and the other used San Francisco. To account for a potential variance in task time depending on city, roughly half the group working on Singlemon first used the Boston task set, and half used the San Francisco task set, and similar for those working with Multimon first. In the second sessions, a participant always used the other city from their experience in the first session. (Ultimately, it turned out that the difference caused by the two task sets was negligible so we collapsed this variable throughout the rest of our analysis). In general, we assumed that participants would perform better in their second session regardless of the condition because of the similarity of tasks across sessions and benefits from familiarity.

3.3 Procedure

Before beginning the experiment, each participant was given a few minutes' training and practice to become familiar with the applications and websites they would be using, if not already so. Participants also completed a pre-study questionnaire that gathered demographic information and data about their experience with Internet tools and multiple monitor computers.

After completing the task set with one display configuration, participants filled out an interim questionnaire. First, they estimated how long it took to complete all the tasks, which we call *perceived task time*. (Later, we will compare the perceived task time and the actual task time; we assumed that if a person estimated the task time as shorter than what it actually was, s/he might have felt less cognitive load than if the opposite was true.) After the time estimation, participants took a NASA TLX survey to assess their subjective workload. The NASA TLX survey measures perceived workload of activities defined by the following factors: mental, physical, and temporal demand; effort;

frustration; and impression of performance [10]. The participant rates the level of each as part of the overall workload. The questionnaire consists of two sections; the first involves making pairwise comparisons between all possible combinations of the above subscales, with the participant choosing one element from each pair that contributed more to overall workload. The second section involves placing a mark on a low-to-high scale that indicates how much of each element contributed to overall workload.

Next, participants moved to the second display configuration and carried out the same process with the alternate task set and city. To conclude, each participant completed a questionnaire that asked them to compare their experiences with the two display settings and we interviewed them about their experiences as well.

We recorded the completion time of each task in the series and we logged window management activities of each session using VibeLog, a window operation logging tool for Microsoft Windows [11]. The recorded actions included opening and closing windows, showing and hiding windows, activating windows, and moving, sizing, minimizing, maximizing, and restoring windows. In addition, we captured a video of their entire screen session using Camtasia Recorder.

3.4 Tasks

We devised a series of typical office productivity and communication-related computing tasks using various MS Office applications, an email reader, an instant messaging (IM) tool, and a web browser to access different travel-related websites. In order to replicate task switching and multitasking in the real world, we designed the scenario to require interleaved use of the different tools. More specifically, participants played the role of an administrative assistant for a company and they were asked to plan a business trip including air travel, hotel stay, and dinner

using the various software tools and websites. The set of tasks included, in order:

- Searching for the lowest roundtrip flight airfare and logging the flight information including departure/arrival time, airline, flight number, and total price to a MS Word file

- Searching for a hotel and saving the reservation information and cost to the MS Word file

- Copying the flight and hotel information to a MS PowerPoint file

- Estimating the total expenses using MS Excel

- Searching for an appropriate restaurant in the city and copying the pertinent information to the PowerPoint file

- Copying directions from the hotel to the restaurant from Internet Explorer to the MS Word file

In addition to performing these main tasks, participants also had to complete side tasks that were independent from the trip planning process. In one side task, the participant had to check email and follow the instructions in a new message. The email included information about an upcoming talk or a book order, depending on the session, and participants were asked to copy that information to a web form at a specified URL. They were asked to check email immediately after copying the hotel information to the PowerPoint file. In the second side task, the participant had to reply to two instant messages. In one message the participant was asked to find a password or a phone number from a specific folder and respond with it, and in the other to check a currency rate or tomorrow's weather forecast on the web. Participants were asked to immediately handle an instant message when it arrived. For each session, instant messages interrupted twice: when participants began copying hotel information from the Word file to the PowerPoint file and when participants began retrieving directions from hotel to the restaurant.

To help maintain similarity of order and performance of steps, participants were given a checklist of the main tasks to perform and were asked to mark each as completed when it was finished.

4 RESULTS AND ANALYSIS

4.1 Multimon vs. Singlemon

We analyzed the task completion times using Repeated Measures Two-way ANOVA, and found a significant main effect of Setting ($F=5.00$, $p=.035$) and Order ($F=114.53$, $p=.000$). Again, Setting indicates whether they interacted with Singlemon or Multimon and Order indicates which setting they used first. To better understand the results, we segmented the measures by session and examined them separately. As Figure 2 shows, participants in the Multimon setting outperformed those in the Singlemon setting in the first session. We also found a learning effect as participants performed faster in their second trial regardless of settings. Though the average task time with Multimon was slightly slower than with Singlemon in the second session, the repeated measures ANOVA showed that Multimon helped reduce task times overall.

Table 1. Task Completion Time

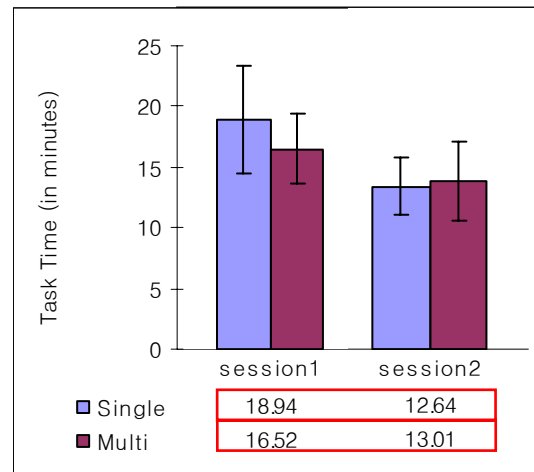|	Setting	Mean	Std. Deviation	N
Session 1	Single	18.94	4.41	13
	Multi	16.52	2.88	14
	Total	17.69	3.83	27
Session 2	Single	12.64	2.27	14
	Multi	13.01	3.11	13
	Total	12.82	2.66	27

Figure 2. Singlemon vs. Multimon in Task Time
Boxes indicate statistically significant differences (p<.05)

Using Repeated Measures Two-way ANOVA again, we analyzed NASA TLX score. We found the main effect of Order ($F=16.796$, $p=.000$) and a trend of Setting ($F=3.181$, $p=.087$). In both sessions, participants felt less workload in Multimon setting than in Singlemon setting, though the amount of reduction was larger in the second session (see Figure 3).

Table 2. NASA TLX Score

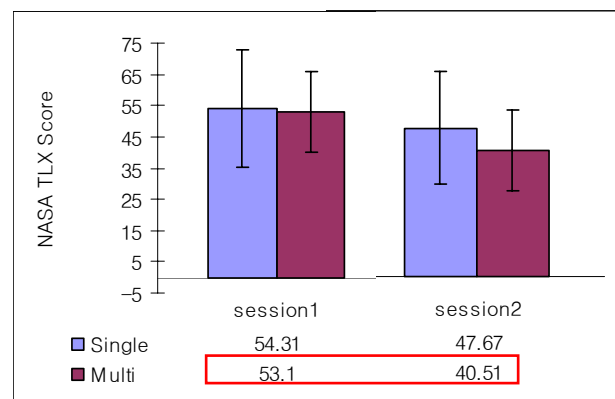|	Setting	Mean	Std. Deviation	N
Session 1	Single	54.31	18.79	13
	Multi	53.10	12.77	14
	Total	53.68	15.65	27
Session 2	Single	47.67	17.93	14
	Multi	40.51	13.13	13
	Total	44.22	15.93	27

Figure 3. Singlemon vs. Multimon in NASA TLX Score

To measure the perceived task times by participants, we analyzed the difference between their estimates of task completion times and actual task completion times. A larger positive number indicates that the person perceived that the task took longer than it did, and may correspond to higher cognitive load. We found no significant differences across settings and orders.

In a post-study questionnaire, participants compared use of the different monitor configurations. Along the dimensions of usefulness, being easy to use, timesaving, and overall impression,

Graphics **Interface** 2008

participants assigned a value where 1 meant Singlemon is much better, 5 meant the two settings are about the same, and 9 meant Multimon is much better. In general, participants strongly favored the multiple monitor setting as shown in Table 3.

Table 3. User Rating on Multimon Compared to Singlemon. Ranges 1 - 9, with 1 meaning Singlemon is better and 9 meaning Multimon is better. 5 indicates even.

		Useful	Easy to use	Timesaving	Overall impression
Single First	Mean	8.38	7.69	7.62	8.23
	N	13	13	13	13
Multi First	Mean	7.64	7.21	5.50	7.36
	N	14	14	14	14
Total	Mean	8.00	7.44	6.52	7.78
	N	27	27	27	27

4.2   Multiple monitor experience

We further analyzed the results to examine the effect of prior experience with multiple monitors. To do this, we used information from the pre-study questionnaire regarding participants' prior experience with Multimon. We then classified each participant into one of two groups: those who are/have been regular multiple monitor users and those who are not.

Both task completion time and NASA TLX score were analyzed using Repeated Measures three-way ANOVA, but we only found the prior main effects of Setting and Order.

Again, for further analysis, we examined the two sessions separately. Figure 4a and 4b illustrate the completion times of both groups in each session. When we examine the results from the first session, prior single monitor-only users performed better with Multimon than with Singlemon. When using the single monitor computer in the first session, multiple monitor users performed better than the participants who regularly use a single monitor computer, though statistically not significant. The results from the second session showed only small differences between two groups and two settings, thus again the learning effect seemed to dominate in the second session. These results were somewhat unexpected since we initially speculated that regular multiple monitor users' performance would improve significantly in the Multimon setting, while performance of those who regularly use a single monitor computer would improve slightly. We will discuss this issue in more detail in the following section.
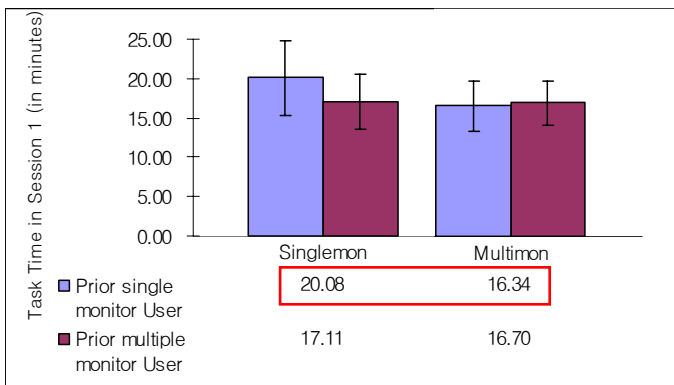

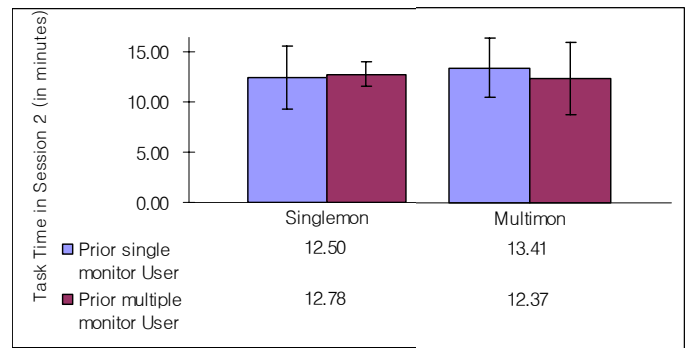Figure 4a. Task Time in Session 1 (in minutes)

Figure 4b. Task Time in Session 2 (in minutes)

We also separated the two groups' workload by session (see Figure 5a and 5b). In the first session experienced multiple monitor users had less workload in both settings, particularly so in the Multimon setting, though statistically not significant. Results from the second session indicated that participants who regularly work with a single monitor had more workload with Singlemon, whereas multiple monitor users felt about the same degree of workload between the two settings. In other words, Multimon appeared to help single monitor computer users reduce workload but did not affect regular multiple monitor users.
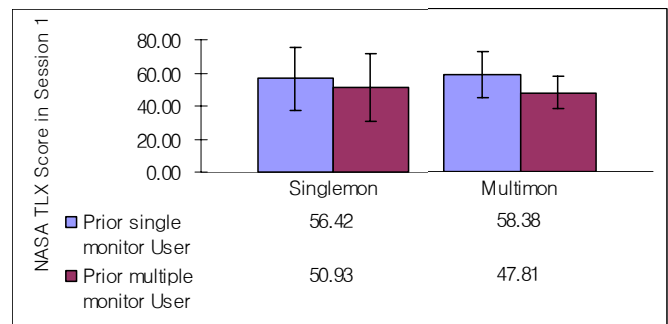

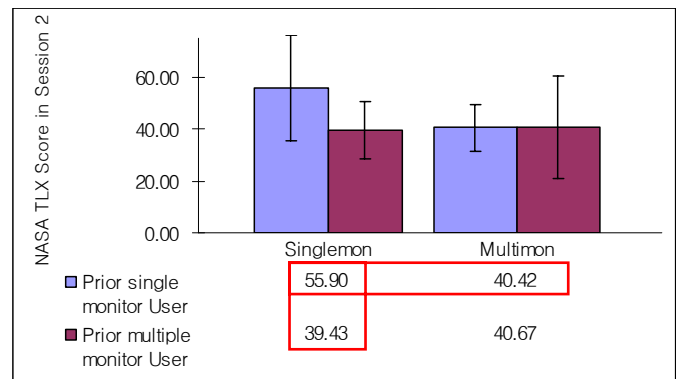Figure 5a. NASA TLX Score in Session 1

Figure 5b. NASA TLX Score in Session 2

While prior experience was not the main effect on task completion time and workload, we found that there was an interaction effect of Setting and Prior Experience (Setting * Prior Experience) in the perceived task time (p=.037). Figure 6a and 6b reveal that in the first session there were only small differences between the actual and perceived task time for both settings. In the second session, however, clear opposite patterns appear between the two user groups. Participants with only single

monitor experience felt that they had performed the tasks slower than they actually had done in the Singlemon setting and faster in the Multimon setting, whereas regular multiple monitor users perceived that they performed slower than actual in the Multimon setting and faster in the Singlemon setting.
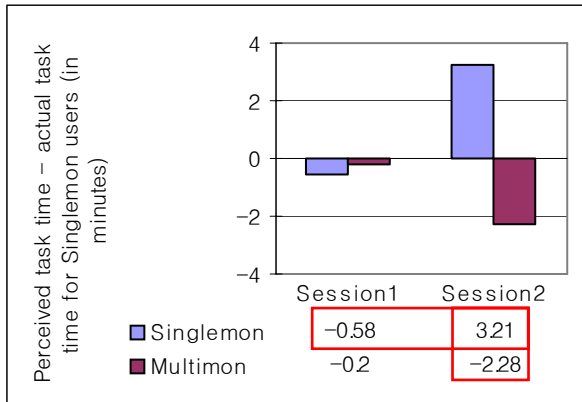

Figure 6a. Singlemon vs. Multimon in (perceived – actual) task time for those who regularly use a single monitor computer
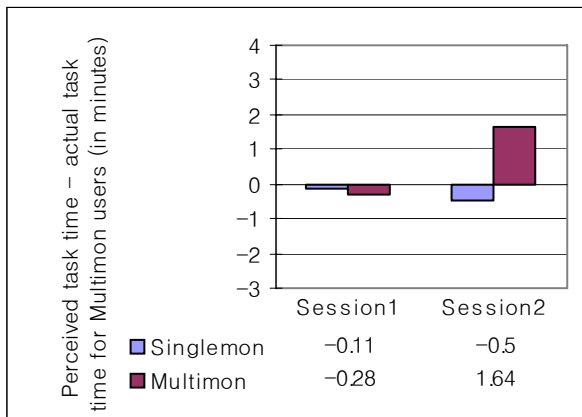
	Session1	Session2
Singlemon	−0.58	3.21
Multimon	−0.2	−2.28

Figure 6b. Singlemon vs. Multimon in (perceived – actual) task time for those who regularly use a multiple monitor computer

	Session1	Session2
Singlemon	−0.11	−0.5
Multimon	−0.28	1.64

There was no main effect of prior experience in terms of subjective preference on the post-study questionnaire.

4.3 Window Management Style

After analyzing the results between experienced and non-experienced multiple monitor users, we were interested if window management style also had an effect on the measures. Since we had logged users' window activities, we were able to examine how they managed multiple applications during the experiment. From the analysis of the logging data, we found two explicit patterns of window management among users: Alt+tab users vs. Move/resizers. Alt+tab users were those who mainly used the Alt+tab function when they needed to switch applications. In contrast, those who we named Move/resizers did not use the Alt+tab function at all or used it infrequently. Instead, they moved and resized the windows to see multiple windows at a time or to traverse between them.

This categorization closely matched that of a previous study of window management practices [13]. That study identified three different ways of organizing screen space: *Maximizers* who used the Alt+tab function to switch windows, *Near Maximizers* who manually resize (nearly) all other windows to occupy all but a little of the remaining portion of the monitor, and *Careful*

Coordinators who tended to have many windows visible simultaneously.

We ran the same analysis as we did previously, but this time we took account into these two user types to identify the effect of window management style. In both task completion time and NASA TLX score, the Repeated Measure 3-way ANOVA results showed that the significant main effects are Setting and Order, and the user type (i.e., Alt+tab users vs. move/resizers) did not significantly influence the results. That is, different window management practices did not significantly affect task time and workload, compared to Setting and Order.

Then we separated results from the two sessions to better understand how the two user types performed differently. Figure 7a shows that both Alt+tab users and Move/resizers benefited from Multimon in the first session, though statistically not significant. In the second session, however, Alt+tab users performed better with Singlemon (see Figure 7b). This might be because the advantage of becoming familiar with task sets outweighed the benefit of using Multimon.
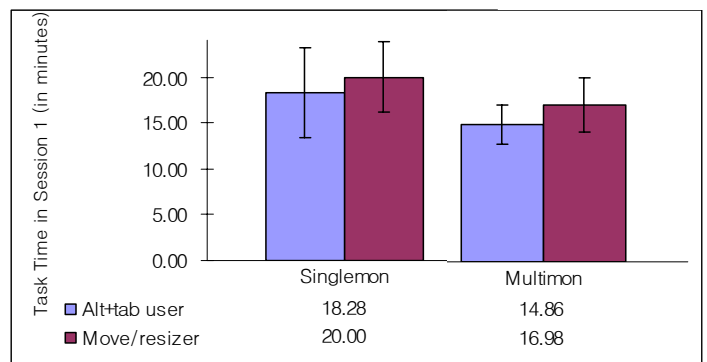

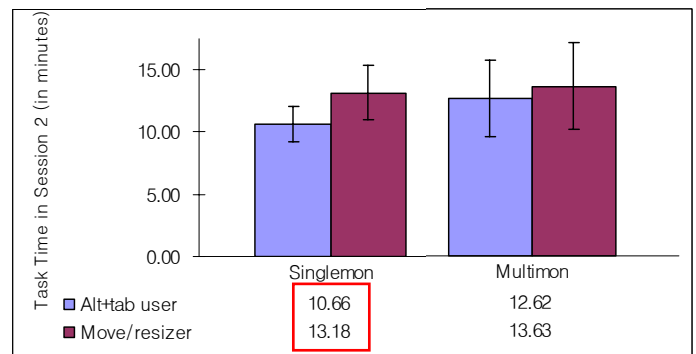Figure 7a. Singlemon vs. Multimon in Task Time for Alt+tab users and Move/resizers. Session 1 Only.

	Singlemon	Multimon
Alt+tab user	18.28	14.86
Move/resizer	20.00	16.98

Figure 7b. Singlemon vs. Multimon in Task Time for Alt+tab users and Move/resizers. Session 2 Only.

	Singlemon	Multimon
Alt+tab user	10.66	12.62
Move/resizer	13.18	13.63

NASA TLX survey results show clear opposite patterns between the two user types. In the first session, Alt+tab users felt much less workload with Multimon while Move/resizers did not benefit from Multimon (see Figure 8a). In the second session, however, Move/resizers felt less workload with Multimon while Alt+tab users benefited more from Singlemon (see Figure 8b).
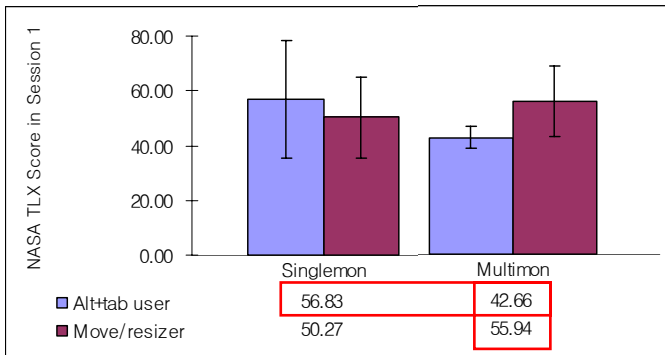
	Singlemon	Multimon
Alt+tab user	56.83	42.66
Move/resizer	50.27	55.94

Figure 8a. Singlemon vs. Multimon in NASA TLX for Alt+tab users and Move/resizers. Session 1 Only.

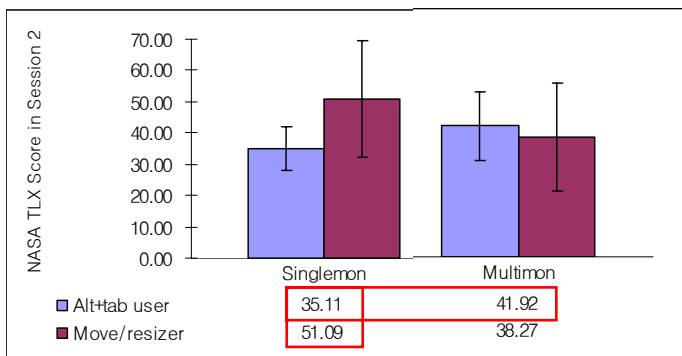|	Singlemon	Multimon
Alt+tab user	35.11	41.92
Move/resizer	51.09	38.27

Figure 8b. Singlemon vs. Multimon in NASA TLX for Alt+tab users and Move/resizers. Session 2 Only.

In regard to the (perceived - actual) task time, we found the main trend of interaction between Setting and User Type (Setting * User Type) (p=.065).

5 DISCUSSION

In general, our analysis showed that the Multimon setting benefited productivity as it resulted in lower task completion time and workload. The results were further echoed by subjective participant ratings on their preference for Multimon. Participants rated Multimon as more useful, easier to use, more timesaving, and had a stronger overall impression.

It should be noted, however, that there was an order effect on the task time. The advantage of Multimon over Singlemon diminished as the users became more familiar with the tasks. This might be because participants knew better how to manage multiple applications for task completion, and they did not have to switch windows as frequently. In the concluding interview, three of the participants noted that they would have found a greater advantage for Multimon if the tasks had been much more complex.

We suspected that the use of multiple monitors might benefit participants differently depending on their prior experience with multiple monitors. People who had been using multiple monitors may benefit more in that they are familiar with such displays and they may have formed their own strategies of use. Also, since they have become accustomed to a Multimon system, using only one monitor might impede their performance. In contrast, users who had little or no experience with multiple monitors likely would benefit less by using the Multimon system. Therefore, we hypothesized that regular multiple monitor users would perform relatively more quickly and feel less work load with the Multimon setting than with the Singlemon setting while a single monitor

users, without prior Multimon computer experience, may show less improvement.

Because the main effect of different user categories was significant only in the (perceived – actual) task time when we considered two sessions together, we examined the two sessions separately and found several results which turned out to be inconsistent with our hypotheses. First, participants who regularly use a single monitor performed better and felt less workload with Multimon than with Singlemon. This indicates that the initial barrier of using Multimon is not as high as we assumed. Even those who first experienced the Multimon setting benefited from using multiple monitors. Second, the reason why those who regularly use multiple monitors did not seem to benefit more by using Multimon was because they performed well with Singlemon as well as Multimon, not because the Singlemon setting impeded their performance. This implies that the two user groups are not mutually exclusive; regular multiple monitor users also use a single monitor setting frequently. From the demographic information of participants, we found that those who described themselves as regular multiple monitor users often use Multimon at a work environment and Singlemon at home. Therefore, using Multimon does not mean moving from Singlemon to Multimon altogether.

We also expected that different window management styles might affect performance under the different settings, though we did not formulate any hypotheses in advance. Again, we examined the two sessions separately and found that in the first session both Alt+tab users and Move/resizers performed faster with Multimon compare to Singlemon. We observed that Alt+tab users did not use or rarely used the Alt+tab function in the Multimon setting across both sessions, and their faster completion in the first trial with Multimon may have occurred because unfamiliarity with the tasks led them to act more like Move/resizers and benefit from the extra space. For Move/resizers, increased screen estate seemed to allow them to coordinate multiple windows more quickly. The logging data showed that the frequency of moving and resizing greatly diminished when they used Multimon. In the second session, however, the benefit of Multimon seemed to deteriorate. The task time of Move/resizers did not differ between Singlemon and Multimon, whereas Alt+tab users performed faster with Singlemon. We speculate that this occurred because of a learning effect. Though participants gained performance benefit from Multimon in the first session, the advantage of becoming familiar with tasks outweighed the advantage of using Multimon in the following trial.

In terms of workload, Alt+tab users had much lower workload with Multimon in the first session, which is consistent with their reduced task time. However, the benefit did not last in the second session. It seemed that once they became comfortable with the tasks, the setting did not influence their cognitive load any more. Actually, participants in the second session who were already familiar with the task set and used the Alt+tab function finished the tasks quite quickly. They did not need to move and resize windows, and switching applications took a very short time. In contrast, Move/resizers felt less workload with Multimon only in the second session. Because they had to manage a larger screen space to allocate multiple windows, the Multimon setting might have imposed greater workload when they first encountered it. As the tasks became familiar, however, they developed a strategy to coordinate multiple windows that allowed them to utilize increased screen space with less workload.

From our observations and interviews, we found that participants' strategy in using Multimon was consistent with Grudin's observations [9], regardless of their prior experience or window management style. The most frequently used strategy was to place one window for main work and another as a reference

window. Participants used the multiple monitors not merely as increased screen space, but as two separate "rooms" where they allocated resources for different purposes.

Finally, we speculate that a number of limitations in the study may have hidden other potential significant findings. First, even though the number of participants suffced to ensure statistical power because we used repeated measures analysis, a larger population may have helped uncover more significant differences. Another issue was the complexity of tasks performed by the participants. We speculate that a more complex task set involving a greater number of applications would show greater benefit for the multiple monitors.

6 CONCLUSION

In this research, we conducted a comparative study to quantitatively examine how people perform common tasks with one or two monitors. Participants performed a business trip planning task using multiple applications and we examined their task completion times, workload measured by NASA TLX, and subjective impressions. We found that multiple monitors helped reduce task time and workload in spite of a learning effect. We also examined how prior multiple monitor usage affected performance and found regular single monitor users benefited from a multiple monitor setting in terms of both task time and workload. Window management patterns also had an effect on performance using multiple monitors. Both Alt+tab users and Move/resizers performed faster with a multiple monitor setting in their first trial. While Alt+tab users benefited more from multiple monitors than a single monitor when tasks were not familiar, Move/resizers benefited more from multiple monitors when they were more familiar with the tasks.

In future work, we would like to further examine the effect of multiple monitors under different contexts. Will the benefit of multiple monitors increase when even more complicated tasks are introduced? How will frequent distractions and peripheral tasks affect usage? We are also interested in examining the learning issues, for example, how will users appropriate the use of multiple monitors over time? When their everyday work requires managing multiple applications and data, how will they develop their own strategy? Answering such questions may require a longitudinal study with users working in a naturalistic setting.

REFERENCES

[1] Ball, R., North, C. Effects of tiled high-resolution display on basic visualization and navigation tasks. *Ext. Abstracts CHI 2005*, ACM Press (2005), 1196-1199.

[2] Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P. Bederson, B., and Zierlinger, A. Drag-and-Pop and Drag-and-Pick: Techniques for Accessing Remote Screen Content on Touch- and Pen-operated Systems. In *Proc. Interact 2003*, IOS Press (2003), 57-64.

[3] Bell, B. and Feiner, S. View Management for Distributed Display Environments, *In Workshop on Distributed Display Environments at CHI 2005*, ACM Press (2005)

[4] Benko, H. and Feiner, S. "Multi-Monitor Mouse". In *Proc. CHI 2005,* ACM Press (2005), 1208-1211.

[5] Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., and Starkweather, G. Toward characterizing the productivity benefits of very large displays. In *Proc. Interact 2003*, IOS Press (2003), 9-16.

[6] De Bruijn, D., De Mul, S., and Van Oostendorp, H. The influence of screen size and text layout on the study of text. *Behaviour and Information Technology, 11,* 2 (1992), 71-78.

[7] Dillon, A., Richardson, J., and McKnight, C. The effects of display size and text splitting on reading lengthy text from screen. *Behaviour and Information Technology, 9*, 3 (1990), 215-227.

[8] Gaylin, K. How are windows used? Some notes on creating empirically-based windowing benchmark task. In *Proc. CHI 1986*, ACM Press (1986), 96-100.

[9] Grudin, J. Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Proc. CHI 2002*, ACM Press (2002), 458-465.

[10] Hart, S. G., and Staveland, L. E. Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.). *Human mental workload*. Elsevier, Amsterdam, 1988, 139–183.

[11] Hutchings, D., Czerwinski, M., Smith, G., Meyers, B., and Robertson, G. Display space usage and window management operation comparisons between single monitor and multiple monitor users. In *Proc. AVI 2004*, ACM Press (2004), 32-39.

[12] Hutchings, D. R. and Stasko, J. mudibo: Multiple Dialog Boxes for Multiple Monitors. In *CHI 2005 Extended Abstracts*, ACM Press (2005), 1471 - 1474.

[13] Hutchings, D. R. and Stasko, J. Revisiting display space management: understanding current practice to inform next-generation design. In *Proc. Graphics Interface 2004*, Canadian Human-Computer Communications Society, 127–134.

[14] St. John, M., Harris, W., & Osga, G. A. (1997). Designing for multitasking environments: Multiple monitors versus multiple windows. In Proceedings of the 41st Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society. pp. 1313-1317.

[15] Kingery D. and Furuta R. Skimming electronic newspaper headlines: a study of typeface, point size, screen resolution and monitor size. *Information Processing and Management 33,* 5 (1997), 685-696.

[16] Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J.A., Verratti, T., and Chiselko, G. Using a large projection screen as an alternative to head-mounted displays for virtual environments. In *Proc. CHI 2000*, ACM Press (2000), 478-485.

[17] Simmons, T. What's the optimum computer display size? *Ergonomics in Design Fall* (2001), 19–25.

[18] Tan, D., Gergle, D., Scupelli, P., and Pausch, R. Similar visual angles, larger display improve spatial performance. In *Proc. of CHI 2003*, ACM Press (2003), 217–224.

[19] Tan, D.S., Gergle, D., Scupelli, P. and Pausch, R. Physically large displays improve path integration in 3D virtual navigation tasks. In *Proc. CHI 2004*, ACM Press (2004), 439-446.

[20] Tan, D.S. and Czerwinski, M. Effects of visual separation and physical continuities when distributing information across multiple displays. In *Proc. INTERACT 2003*, IOS Press (2003), 252-265.