# Productivity and Multi-Screen Computer Displays

*Rocky Mountain Communication Review*
*Volume 2:1, Summer, 2004*
*Pages 31-53*

Janet Colvin
Nancy Tobler
James A. Anderson

*One hundred eight university and non university personnel participated in a comparison of single monitor, multi-monitor and multi-monitor with Hydravision display configurations. Respondents edited slide shows, spreadsheets and text documents in a simulation of office work, using each of the display arrays. Performance measures, including task time, editing time, number of edits completed, and number of errors made and usability measures evaluating effectiveness, comfort, learning ease, time to productivity, quickness of recovery from mistakes, ease of task tracking, ability to maintain task focus and ease of movement among sources were combined into an overall evaluation of productivity. Multi-screens scored significantly higher on every measure.*

As processor speed and memory capacity have increased and become less expensive, the office has found that it can support more open applications and that multi-tasking could be a reality, not just a term. One problem remaining has been the management of the computer desktop. Even with increased monitor size, the single screen presents fundamental problems with window placement, stacking and tracking windows, multiple applications on the task bar, and the like (Delefino, 1993; Grudin, 2003). These problems have limited the increases in productivity theoretically made possible with increased processor speed and memory capacity.

## Multi-Screen Solution

Solutions to this problem have been available since the advent of the Windows 98 operating system that allows the PC platform to support multi-monitor displays. Initially, multi-screen configurations found use in computer gaming but has found little interest or recognition in the business or academic community. Part of the reason for that lack of interest has been the absence of evidence of value (Binder, 2001; Lindsley, 1996; St. John, Harris, & Osga, 1997). This study addresses that absence by comparing multi- and single-screen configurations across performance and usability measures.

## Multi-Screen Configurations

The multi-screen display configuration can range from a fully integrated set of liquid crystal displays to a simple, physical arrangement of two or more CRT monitors (Bohannon, 2003; Dyson, 2002; Vellotti, 2001). Each screen or monitor in a multi-screen display is connected to the same computer through its own display port and is treated by the operating system both as a unified, boundaried space and as a connected or extended desktop. For example, an application will maximize to the boundaries of its “home” single screen but can also be “windowed” across all screens (a number theoretically unlimited but usually 2-5). Multi-screens configurations allow the user to place different windows on different screens or to spread a single application across all available screens (Brown & Ruf, 1989).

## Multi-Screen Management Software

Multi-screen management software adds another

*Janet Colvin and Nancy Tobler are doctoral candidates and James A. Anderson is Professor and Director of Graduate Studies in the Department of Communication at the University of Utah, Salt Lake City, UT 84112..*

potential set of efficiencies (Randall, 1999). Multi-screen management software allows the user to instantly transport application windows to different screens, maximize applications across all displays, open child windows (e.g., multiple spreadsheets or tool and property sub-menus) on different displays, and to switch between virtual desktops (e.g., from a text editing set up to a graphics design set up).

### Productivity and Multi-screen Displays

Productivity testing involves the reproduction of an ordinary work site, plausible and recognizable work tasks, and reasonable conditions of work (Sherry & Wilson, 1996; Stolovich & Keeps, 1992). Productivity testing then is a combination of performance testing and usability testing (Baurua, Kriebel, & Mukhopadhyay, 1991; Jorgenson & Stiroh, 1999). In performance testing, automated tools collect facts about what the users *actually did* and *how long it took* them to do it. In usability testing, a sample group is asked to perform a set of tasks and *subjectively rate the ease of use* of a piece of hardware or software. Because usability without increased performance or increased performance without adequate usability will not sustain overall increases in productivity, authentic measures of productivity must involve both (Brynjolfsson & Yang, 1996).

## Comparing Single and Multi-Screens over Performance and Usability

### Overview

In order to test the productivity of multi-screen configurations, an experimental comparison was devised using three blocks of simulated office tasks. Each block contained a text editing task (TXT), a spreadsheet editing task (SST) and a slide presentation editing task (PPT). Each task within the block was designed to use six windows of information: Two windows concerned the administrative, data collection, and simulation management of the experiment per se and four windows were components of the task. A seventh window provided navigational information that governed the entire session and contained the hyperlinks for the various files required.

Each of the 108 respondents completed one different block in each of three configurations: single screen (SS), multi-screen (MS), and multi-screen assisted by multi-screen management software (HV)[1]. The order of tasks was the same in each block: text, spreadsheet and slide. An equal number of respondents (36 per block x configuration combinations) completed each block to control for possible task by configuration differences. Screen configurations and tasks were used as "within subjects" factors in the analysis.

Strong order effects were to be expected as respondents learned how the task was to be performed. To control for these effects, an equal number of respondents (12 per each of the 9 block x configuration x order combinations) started the task set with a different configuration in the first position. Table 1 presents the rotation of tasks and configurations. This procedure was repeated for each of the task sets. Order effects were, therefore, balanced across all configurations. In this manner, each respondent completed all 9 tasks in blocks of three and experienced all three screen configurations addressing them in one of three orders.

Finally, to get some sense of an "optimal" number of monitors, the multi-screen configuration was further divided into one with two monitors and one with three monitors. Half of the respondent pool (54) worked the tasks in a 2-monitor setup and half in a 3-monitor setup. This "monitor condition" was used as a "between-subject" factor in the analysis.

### Tasks

All three tasks were based on the same scenario: A destination text, spreadsheet, or slide presentation had been previously prepared and sent out for review or error correction. The copy edits and corrections

**Table 1**

| **Order** | | | **First Task** | | | **Second Task** | | | **Third Task** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Start** | **#Rs** | Text | Spread | Slide | Text | Spread | Slide | Text | Spread | Slide |
| **Single** | 12 | GS | CR | MDY | SR | CS | WP | HV | PR | MDK |
| **Single** | 12 | HV | PR | MDK | GS | CR | MDY | SR | CS | WP |
| **Single** | 12 | SR | CS | WP | HV | PR | MDK | GS | CR | MDY |
| **Multi** | 12 | GS | CR | MDY | SR | CS | WP | HV | PR | MDK |
| **Multi** | 12 | HV | PR | MDK | GS | CR | MDY | SR | CS | WP |
| **Multi** | 12 | SR | CS | WP | HV | PR | MDK | GS | CR | MDY |
| **HV** | 12 | GS | CR | MDY | SR | CS | WP | HV | PR | MDK |
| **HV** | 12 | HV | PR | MDK | GS | CR | MDY | SR | CS | WP |
| **HV** | 12 | SR | CS | WP | HV | PR | MDK | GS | CR | MDY |

Text tasks: Graduate studies, Screen Report, Hydravision
Spreadsheet tasks: Candidate Rankings, Products by Region, Customer Survey
Slide Tasks: Multi-Desk, Multi- Display, Window Placement

**Table 1: Starting rotation of tasks and configurations.**

had been returned to the respondent whose job was now to make the changes on the destination file.

*Text Tasks*

The text files were prepared using Microsoft Word© with "track changes" enabled. The task files consisted of the destination document on which all changes were to be made and two source documents (called Mulcahy Edit and Tobler Edit) from which the changes were to be drawn. Each of the source documents had between 8 and 10 edits to be completed, including a requirement to open a graphics file and to copy and paste a graphic. (A full report of this study is available at http://www.necmitsubishi.com/solutions/SolutionDetail.cfm?solution=293&Document=1138). The three texts were well populated with position markers such as paragraphing, headings, and graphics to assist the respondent in tracking locations from one document to another.

*Spreadsheet Tasks*

The spreadsheet files were prepared using Microsoft Excel© and Microsoft Word©. Each spreadsheet was designed to cover approximately one and a half screens (an average of 33 rows by 25 columns). Each of the data sets had summary information that was dynamically linked to a bar chart. Corrections were provided to the respondents in the form of a "Corrections Memo" simulating an e-mail addressed to them. Sixteen corrections were listed for the respondent to enter. After the corrections were made, the respondent was to copy the bar chart and paste it into a designated location in a "Final Report." The Final Report was accessed by a hyperlink on the instructions page.

*Slide Tasks*

The slide files were prepared using Microsoft PowerPoint©. PowerPoint has a rather limited editing handling protocol (as compared with most word processing). Edits were identified in comments and placed in the source documents in color-coded type. Each slide task had between 11 and 17 edits, including navigating to a graphics page, selecting a logo, and pasting the graphic into a new slide.

**Data Collection**

Data were collected in six ways: a paper and pencil intake questionnaire, automated time reports and automated usability questionnaire, stop watch measurements, task observations, and open-ended, end-of-testing questions. A description of each follows:

*Intake Questionnaire*

A single page intake questionnaire asked respondents to record their experience levels with computers, with the various applications used in the study, and with multiple screens. It also queried job experience and hours of work.

*Time Report and Usability Questionnaire*

An Excel spreadsheet was devised to collect the respondent's ID number, the time spent reading instructions, the total time spent on the task, and the responses to each of 8 usability questions. The usability questionnaire recorded the respondents self reports on their effectiveness, comfort, ease of learning, productivity, mistake recovery, task tracking, task focus, and ease of movement across sources (adapted from Lewis, 1995; Davis, 1989).

*Stop Watch Measurements*

Stop watch timing was initiated at the start of the actual editing task. Each task had its own marker events for the start and completion of editing. An observer/facilitator (O/F), seated next to the respondent, started the watch on the initiation event and stopped it on the completion event. The time values were recorded on the task observation sheet in minutes and seconds.

*Task Observations*

As an observer (facilitation practices are described under "Protocol"), the O/F was responsible for stop watch data, recording the correct completion of each edit, recording any missed edits and errors in editing or changes otherwise introduced into the source documents, recording any comments about the task or the screen configuration, and any unusual practices in the editing task that appeared worthy of notice.

*Post-Session Questions*

At the completion of all the tasks, the O/F asked four questions: "Focusing on single screen versus multiple screens, what did you think about that difference?" "Focusing on multiple screens with Hydravision and multiple screens without Hydravision, what did you think about that difference?" "Focusing on the tasks and the different screen configurations, did any task seem easier or harder in a given screen configuration?" "Focusing on the experiment itself, was there anything that bothered you or that we should do differently?" A summary of the respondent's answers was recorded on the task observation sheet.

## Protocol

### *Sampling*

Using a combination of advertisements and snowball sampling, 108 respondents were drawn from students, staff, and faculty from the university and individuals from the larger community as well. The sample was equally divided between the 2-monitor and 3-monitor conditions.

### *Testing Procedures*

Upon arrival, the respondent was given a short description of the study and the intake questionnaire to complete. The respondent was then shown one of three 5-minute training videos, SS, MS, or HV depending on the initial configuration of the task. The training video demonstrated a set of editing procedures appropriate to each task in the block and to the specific screen configuration.

At the conclusion of the video, the O/F described the screen configuration that was in use, the tasks to be done, and the role the O/F would play in the process. When all questions were answered, the respondent was asked to navigate to the first time stamp screen to begin the block session. When the respondent initiated the editing task, the stopwatch was started. Respondents were given 5 minutes to complete the task, although time was added to allow the completion of an edit in progress. The O/F recorded each edit as it was made. Errors and missed edits were also recorded. At the conclusion of the task, the stopwatch was stopped, the time recorded, and the respondent immediately directed back to the time stamp. The respondent then checked the task "Done Box," completed the usability questionnaire, and posted the file. This procedure was repeated for each configuration. At the conclusion of all three task blocks, post-session questions were asked and answers recorded. Each session took approximately 90 minutes. Respondents were paid $20 for their time.

Project activities were under the supervision of a project ethicist whose responsibility was to ensure that all procedures were followed by the O/F and other project staff. The project ethicist made random visitations and observed entire sessions. Her final report noted no violations.

### *Facilities*

Testing was done in the University of Utah, Department of Communication interaction laboratory. This testing facility has the look and feel of a living room (albeit one with a large one-way mirror and video cameras in the corners) with couches, easy chairs and a large television set. Two work tables were added for each of the testing stations.

Each testing station was configured with a new PC computer with a clean install of Windows XP and Microsoft Office Suite. The computers were based on the Intel Pentium 4 chip running at 1.8 GHz, with 512MB DDR SDRAM, a 40 GB, 7200rpm Ultra ATA hard drive, standard keyboard, and two button wheel mouse. Monitors were NEC Mitsubishi Multisync LCD 1855NX, an 18 inch liquid crystal display. Display boards were ATI Radeon 9000 AGP with two monitor ports and ATI Radeon 7000 PCI with a single port. One station had two monitors arranged in a slight V with the right hand monitor having the taskbar; the other had three monitors in a triptych arrangement with the task bar on the center monitor.

## Performance Measures

### *Basic Variables and Their Definitions*

Five variables used to test performance were collected automatically or through direct observation. The variables, their definitions and method of collection are reported below:

*Task Time:* One of two basic time units. Task Time is the lapsed time from the respondent's checking of the task "Start" Box on the Time Stamp to the Respondent's checking of the "Done" box on the Time Stamp. Task Time includes set up time and edit time plus any time spent in meeting project requirements (navigating to and from the Time Stamp, for example). Task Time was an automated data collection.

*Edit Time:* The other basic time unit. Edit Time is the stopwatch recorded time from the first editing marker event to the last editing marker event. It represents the amount of time actually on task and has no other time component. The time was recorded by the O/F assigned to the respondent.

*Number of Correct Edits:* The number of correctly executed edits as observed and recorded by the O/F. Each of these edits were listed for each task on the task observation sheets. The O/F checked off each edit as it was completed or recorded an error or a miss as described below.

*Number of Errors:* The O/F recorded an error when the edit called for was completed incorrectly. An error was defined as any event that would have to be "found" and "corrected" by another editing process.

*Number of Missed Edits:* The O/F recorded a missed edit when the respondent skipped a complete edit (partial edits were considered errors).

**Derived Variables and Their Definitions**

Five performance variables were derived through calculations using the basic variables as factors. Those variables and their definitions are:

*Proportion of Edits Completed:* The number of correct edits divided by the total number of edits required by the task.

*Accuracy:* The number of correct edits minus the number of errors and missed edits. Accuracy is a performance cost measure. Inaccurate editing increases costs as the task has to be redone. The greater the inaccuracy, the less confidence can be given to the original work and the more care required in the re-editing.

*Proportion of Accurate Edits:* The accuracy coefficient (number of correct edits minus the number of errors and missed edits) divided by the number of edits required.

*Time per Edit:* Edit time divided by the number of correct edits. This measure can be used to project the time required for larger tasks.

*Time per Accurate Edit:* Edit time divided by the accuracy coefficient (number of correct edits minus the number of errors and missed edits).

**Questionnaires**

Two questionnaire instruments were used in this study: an intake questionnaire that queried respondents on their computer, application, and multi-screen experience and a usability questionnaire administered after every task performance. An open-ended interview based on four questions followed the testing session.

*Intake Questionnaire*

The intake questionnaire was a paper and pencil device composed of 6 sections: *Computer Expertise:* A four point scale ranging from zero (None) to 3 (Advanced). *Application Expertise:* A four point scale ranging from zero (None) to 3 (Advanced). *Block Expertise:* The average of the three Application Expertise measures divided into three roughly equal groups. Cut points were (1) less than 1.67, (2) equal to 1.67, and (3) 2.00 or greater.

Corrections were made for anomalous cases (described in the Performance by Expertise section). *Time Spent on Text, Spreadsheet, and Slide Applications:* In hours per week from zero to ten. *Level of Application Use:* A three point scale from one (Personal) to three (Professional). *Multi-screen Experience:* A "yes" "no" item followed by the number of monitors used (1-6). *Current Job Situation:* Number of hours per week on the job and job title. Data were hand entered with double entry verification.

*Usability Questionnaire*

Each task performance was immediately followed by a usability questionnaire that was part of the time stamp file. The questionnaire used a 10-point slider to register the self-reported position between the poles of Strongly Disagree and Strongly Agree. As reported above, the items recorded the respondents' self reports on their effectiveness, comfort, ease of learning, productivity, mistake recovery, task tracking, task focus, and ease of movement across sources. Data were recorded by the same procedures used in collecting the time data and directly entered into the data base.

*Interviews*

Respondents were asked to compare single and multi-screens, multi-screens with software and without, task difficulty in different configurations, and to comment on the protocol itself. A summary of each response was recorded by the O/F and entered verbatim into the data file.

## Analysis and Results: Performance Data

### Statistical Design: Task Variables

In each of the 12 basic and derived variables, data were reorganized from their original task-specific entry into a task-type centered entry that distributed both order of performance and specific task in balanced numbers throughout the data. Each task-type data set had an equal number of the three tasks and three orders.

All respondents did all task-types in all screen configurations (a different version of the task type was used in each configuration). All performance variables are, therefore, "within subjects" or "repeated measures" variables. This design controlled for inter-subject differences. The two "within" variables in this design, then, were task types (*Tasks*) and configurations (*Screens*). The task types were *Slide*, *Spreadsheet*, and *Text*. The three configurations were single screen (*SS*) multi-screen (*MS*) and multi-screen with *Hydravision* (*HV*). The testing condition of a two-monitor or three-monitor station was a "between subjects" or "independent groups" factor in the design. Half of the respondents went through the protocol in each of these *conditions*. Figure 1

**Figure 1**

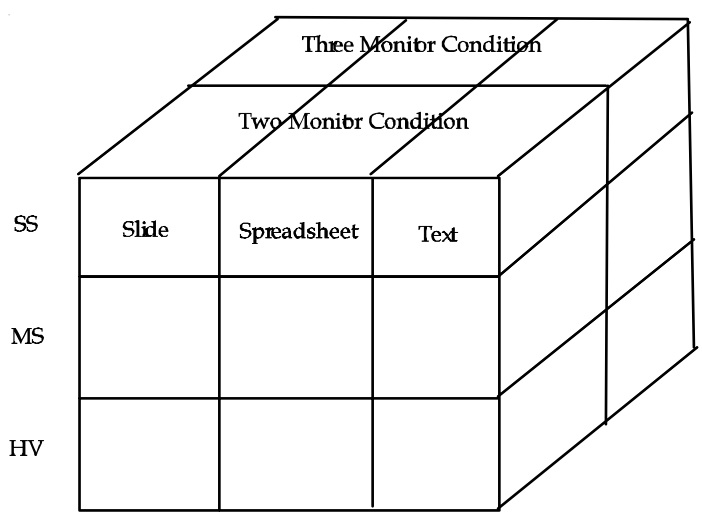


**Figure 1: Statistical design for each performance variable, tasks by screens by conditions.**

diagrams the design.

Each task variable was analyzed using this classic "Type III" design using the General Linear Model as formulated in SPSS. An Alpha of .05 was set as the decision criterion for significance.

**Performance Results**

The restricted space of this venue allows us to present the results for only four of the 12 variables—Edit Time, Number of Correct Edits, Accuracy, and Time per Edit. There is a great deal of redundancy in the 12 variable set; these four variables were selected because they show the greatest amount of unique information. Each report starts with tests of significance in the three-factor (tasks by screens by conditions), each of the two-factor (screens by tasks, screens by conditions, tasks by conditions) and main effects (screens, tasks, and conditions). A table of the means, standard deviations, and confidence intervals by cell is then presented followed by tables of means and standard deviations for each significant condition. The reader is reminded that significant interactions at one level confound the analysis of the next lower level (three-factor confounds two-factor confounds main effects). The results will be discussed only to the lowest non-confounded level.

**Table 2**

| Factors | F-Test | Degrees of Freedom | Significance |
|---|---|---|---|
| Screens by Tasks by Condition | 2.732 | 4, 424 | .029 |
| Screens by Condition | .013 | 2, 212 | .987 |
| Tasks by Condition | .643 | 2, 212 | .527 |
| Screens by Task | 9.915 | 4, 424 | .000 |
| Screens | 21.254 | 2, 212 | .000 |
| Tasks | 200.681 | 2, 212 | .000 |
| Conditions | .026 | 1, 106 | .872 |

**Table 2: Analysis of variance results for Edit Time.**

*Edit Time*

The Edit Time variable measured the time lapsed between the first task marker event and the last task marker event. It can be considered as on-task time. Table 2 presents the analysis of variance results. The ANOVA indicates a three-way interaction between screens, tasks, and conditions. Two-way interaction effects were checked. The screens by task F-test indicates an interaction effect. These findings demonstrate that tasks (slides, spreadsheet, and text) are not consistent over time. Table 3 presents the means, standard deviations, and confidence intervals for tasks and screens by condition for the edit time variable. Table 4 presents a comparison of the SS means with the MS means for the two-monitor conditions. Table 5 presents a comparison of the SS means with HV means for the two- and three-monitor conditions. Significant differences are noted in these latter tables.

**Table 3**

| Monitors | Screens | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| Two Monitors | Single | Slide | 275.444 | 6.165 | 263.221 | 287.668 |
| | | Spreadsheet | 243.037 | 7.326 | 228.512 | 257.562 |
| | | Text | 297.481 | 3.724 | 290.099 | 304.864 |

**Table 3 (Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| Two Monitors | Multi-screen | Slide | 266.796 | 6.148 | 254.608 | 278.984 |
| | | Spreadsheet | 213.056 | 7.721 | 197.747 | 228.364 |
| | | Text | 279.759 | 6.467 | 266.937 | 292.581 |
| | Hydravison | Slide | 261.611 | 7.112 | 247.511 | 275.712 |
| | | Spreadsheet | 217.648 | 7.361 | 203.054 | 232.243 |
| | | Text | 280.352 | 5.006 | 270.426 | 290.277 |
| Three Monitors | Single | Slide | 272.833 | 6.165 | 260.610 | 285.057 |
| | | Spreadsheet | 253.389 | 7.326 | 238.864 | 267.914 |
| | | Text | 293.648 | 3.724 | 286.265 | 301.031 |
| | Multi-screen | Slide | 275.148 | 6.148 | 262.960 | 287.336 |
| | | Spreadsheet | 209.296 | 7.721 | 193.988 | 224.605 |
| | | Text | 279.074 | 6.467 | 266.252 | 291.896 |
| | Hydravision | Slide | 270.889 | 7.112 | 256.788 | 284.989 |
| | | Spreadsheet | 204.389 | 7.361 | 189.794 | 218.983 |
| | | Text | 285.389 | 5.006 | 275.463 | 295.315 |

**Table 3: Conditions by screen configurations by tasks means, standard errors, and confidence intervals for Edit Time.**

**Table 4**

| | Task | Single Mean | Multi Mean | Difference | Percent Change | Significant |
|---|---|---|---|---|---|---|
| Two Monitors | Slide | 275.444 | 266.796 | 8.648 | 3 | No |
| | Spreadsheet | 243.037 | 213.056 | 29.981 | 12 | Yes |
| | Text | 297.481 | 279.759 | 17.722 | 6 | Yes |
| Three Monitors | Slide | 272.833 | 275.148 | -2.315 | -1 | No |
| | Spreadsheet | 253.389 | 209.296 | 44.093 | 17 | Yes |
| | Text | 293.648 | 279.074 | 14.574 | 5 | Yes |

**Table 4: Comparison of SS screen Edit Time means with MS Edit Time means, difference, percent of change, and significance for each monitor condition.**

**Table 5**

| | Task | Single Mean | Hydravision | Difference | Percent Change | Significant |
|---|---|---|---|---|---|---|
| Two Monitors | Slide | 275.444 | 261.611 | 13.833 | 5 | No |
| | Spreadsheet | 243.037 | 217.648 | 25.389 | 10 | Yes |
| | Text | 297.481 | 280.352 | 17.129 | 6 | Yes |
| Three Monitors | Slide | 272.833 | 270.889 | 1.944 | 1 | No |
| | Spreadsheet | 253.389 | 204.389 | 49.00 | 20 | Yes |
| | Text | 293.648 | 285.389 | 8.259 | 3 | Yes |

**Table 5: Comparison of SS screen Edit Time means with HV Edit Time means, difference, percent of change, and significance for each monitor condition.**

The significant three-factor interaction requires analysis at the cell level. Eleven of the 12 comparisons between single screen and multi-screen configurations showed reductions in editing time. These differences were significant in eight of these comparisons. Only the slide task failed to show significant or consistent reductions with the three-monitor MS condition showing a reversal. The spreadsheet tasks showed the largest reductions of time across both monitor conditions. Slide and text editing was done more quickly in the two-monitor condition; spreadsheet editing was faster in the three-monitor condition. None of the differences were significant, although nearly so in the spreadsheet task.

*Number of Edits*

The "Number of Edits" variable gave a count of the number of edits correctly entered by the respondent. This measure is a typical measure of productivity (number of units produced). Table 6 presents the analysis of variance results.

The results in Table 6 indicate that the differences among screen configurations changed over tasks. The lack of a significant three-factor interaction or any of the two-factor interactions involving the condition of two or three monitors signals that the configuration means and the task means remained consistent over the monitor conditions. Although no interaction involving condition was significant, Table 7 presents the cell means and the single screen multi-screen comparisons and Tables 8 and 9 present the comparisons of single screen to multi-screen and single screen to Hydravision to keep the data record consistent for the reader. Tables 10 through 12 present the break down of the significant screens by task interaction.

In Tables 10 through 12, multi-screen configurations show a consistent increase in the number of edits completed over single screen. This advantage is significant in five of the six comparisons. In the lone non-significant condition, the MS mean is .02 below the upper bound limit of the SS confidence interval and is matched with a significant difference in the HV multi-screen condition. It is likely that the multi-

**Table 6**

| Factors | F-Test | Degrees of Freedom | Significance |
|---|---|---|---|
| Screens by Tasks by Condition | 1.330 | 4, 424 | .258 |
| Screens by Condition | .192 | 2, 212 | .826 |
| Tasks by Condition | .390 | 2, 212 | .678 |
| Screens by Task | 4.796 | 4, 424 | .001 |
| Screens | 25.541 | 2, 212 | .000 |
| Tasks | 585.306 | 2, 212 | .000 |
| Conditions | .287 | 1, 106 | .594 |

**Table 6: Analysis of variance results for Number of Edits.**

**Table 7**

| Monitors | Screens | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| Two Monitors | Single | Slide | 10.741 | .446 | 9.856 | 11.626 |
| | | Spreadsheet | 17.148 | .395 | 16.365 | 17.932 |
| | | Text | 11.389 | .464 | 10.469 | 12.309 |

**Table 7 (Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| Two Monitors | Multi-screen | Slide | 11.852 | .492 | 10.877 | 12.827 |
| | | Spreadsheet | 17.630 | .285 | 17.064 | 18.195 |
| | | Text | 13.500 | .442 | 12.623 | 14.377 |
| | Hydravison | Slide | 11.741 | .435 | 10.878 | 12.603 |
| | | Spreadsheet | 17.722 | .353 | 17.022 | 18.422 |
| | | Text | 14.037 | .384 | 13.277 | 14.798 |
| Three Monitors | Single | Slide | 11.037 | .446 | 10.152 | 11.922 |
| | | Spreadsheet | 16.852 | .395 | 16.068 | 17.635 |
| | | Text | 11.796 | .464 | 10.876 | 12.717 |
| | Multi-screen | Slide | 11.130 | .492 | 10.154 | 12.105 |
| | | Spreadsheet | 18.278 | .285 | 17.712 | 18.843 |
| | | Text | 13.852 | .442 | 12.975 | 14.729 |
| | Hydravision | Slide | 12.259 | .435 | 11.397 | 13.122 |
| | | Spreadsheet | 17.944 | .353 | 17.244 | 18.644 |
| | | Text | 14.352 | .384 | 13.591 | 15.112 |

**Table 7: Conditions by screen configurations by tasks means and standard errors for Number of Edits.**

**Table 8**

| | Task | Single Mean | Multi Mean | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 10.741 | 11.852 | 1.111 | 10 |
| | Spreadsheet | 17.148 | 17.630 | 0.482 | 2 |
| | Text | 11.389 | 13.500 | 2.111 | 18 |
| Three Monitors | Slide | 11.037 | 11.130 | 0.093 | 1 |
| | Spreadsheet | 16.852 | 18.278 | 1.426 | 8 |
| | Text | 11.796 | 13.852 | 2.056 | 17 |

**Table 8: Comparison of SS screen Number of Edits means with MS Number of Edits means, difference, percent of change for each monitor condition.**

**Table 9**

| | Task | Single Mean | Hydravision | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 10.741 | 11.741 | 1.00 | 9 |
| | Spreadsheet | 17.148 | 17.722 | 0.574 | 3 |
| | Text | 11.389 | 14.037 | 2.648 | 23 |
| Three Monitors | Slide | 11.037 | 12.259 | 1.222 | 11 |
| | Spreadsheet | 16.852 | 17.944 | 1.092 | 6 |
| | Text | 11.796 | 14.352 | 2.556 | 21 |

**Table 9: Comparison of SS screen Number of Edits means with HV Number of Edits means, difference, percent of change for each monitor condition.**

**Table 10**

| Configuration | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| Single | Slide | 10.889 | .316 | 10.263 | 11.515 |
| | Spreadsheet | 17.000 | .279 | 16.446 | 17.554 |
| | Text | 11.593 | .328 | 10.942 | 12.243 |
| Multi-screen | Slide | 11.491 | .348 | 10.801 | 12.180 |
| | Spreadsheet | 17.954 | .202 | 17.554 | 18.353 |
| | Text | 13.676 | .313 | 13.056 | 14.296 |
| Hydravison | Slide | 12.000 | .308 | 11.390 | 12.610 |
| | Spreadsheet | 17.833 | .250 | 17.338 | 18.328 |
| | Text | 14.194 | .271 | 13.657 | 14.732 |

**Table 10: Screen configurations by tasks means, standard errors, and confidence intervals for Number of Edits.**

**Table 11**

| Task | Single Mean | Multi-screen | Difference | Percent Change | Significance |
|---|---|---|---|---|---|
| Slide | 10.889 | 11.491 | 0.602 | 6 | No ($\alpha$ = .15) |
| Spreadsheet | 17.00 | 17.954 | 0.954 | 6 | Yes |
| Text | 11.593 | 13.676 | 2.083 | 18 | Yes |

**Table 11: Comparison of SS screen Number of Edits means with MS Number of Edits means, difference, percent of change, and significance.**

**Table 12**

| Task | Single Mean | Hydravision | Difference | Percent Change | Significance |
|---|---|---|---|---|---|
| Slide | 10.889 | 12 | 1.111 | 10 | Yes |
| Spreadsheet | 17.00 | 17.833 | 0.833 | 5 | Yes |
| Text | 11.593 | 14.194 | 2.601 | 22 | Yes |

**Table 12: Comparison of SS screen Number of Edits means with HV Number of Edits means, difference, percent of change, and significance.**

screen advantage in this case is being masked by sample conditions. In looking at screen by task effects, the limited experience of slide editing minimizes the overall effect.

*Accuracy*

Accuracy is a constructed variable based on the number of completed edits minus the number of error and the number of misses. The rationale for this measure is that missed work and incorrect work requires more time and money to correct than simple unfinished work. While an analysis of edits and errors indicated an advantage for multi-screen configurations, it is possible that the location of these measures may result in a different outcome. That possibility suggests that should the same advantage appear in Accuracy, it is a confirmation rather than a replication. Table 13 presents the analysis of variance results.

The three-factor interaction and the two-factor interactions involving the number of monitors were not significant, but the two-factor screens by task interaction was. Table 14 presents the means, standard errors, and confidence intervals for the cell values; Table 15 presents a comparison of SS and MS means; Table 16 presents a comparison of SS and HV means, all for the data record.

Because the three-factor interaction was not significant and the two-factor screens by task interaction was, the data are best analyzed by collapsing monitor conditions and looking at the screens means by task. Table 17 presents that information. Inspection of Table 17 shows that multi-screen configurations resulted in higher accuracy scores that were significantly higher in all but the SS to MS slide task compaison ($\alpha$ = .125). In addition, the HV text scores were significantly higher than the MS text score, although the other two comparisons were not significant and their direction mixed.

*Time per Completed Edit*

Time per Completed Edit is the editing time divided by the number of completed edits. It represents the flow of work over time and can be used to craft estimates of work completion over jobs of varying length. Table 18 presents the analysis of variance results for Time per Completed Edit. None of the multi-factor interactions involving Condition were significant. The two-factor Screens by Tasks interaction was significant, pointing to a differential effect of screen configurations across tasks. Tables

**Table 13**

| Factors | F-Test | Degrees of Freedom | Significance |
|---|---|---|---|
| Screens by Tasks by Condition | 2.065 | 4, 424 | .085 |
| Screens by Condition | .026 | 2, 212 | .974 |
| Tasks by Condition | 3.028 | 2, 212 | .697 |
| Screens by Task | 3.850 | 4, 424 | .004 |
| Screens | 22.610 | 2, 212 | .000 |
| Tasks | 357.961 | 2, 212 | .000 |
| Conditions | .410 | 1, 106 | .523 |

**Table 13: Analysis of variance results for Accuracy.**

**Table 14**

| Monitors | Screens | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| Two Monitors | Single | Slide | 10.130 | .483 | 9.173 | 11.087 |
| | | Spreadsheet | 16.315 | .528 | 15.269 | 17.361 |
| | | Text | 10.500 | .502 | 9.506 | 11.494 |
| | Multi-screen | Slide | 11.315 | .514 | 10.296 | 12.333 |
| | | Spreadsheet | 16.630 | .498 | 15.643 | 17.617 |
| | | Text | 12.796 | .497 | 11.810 | 13.782 |

**Table 14 (Continued)**

| | | | | | | |
|---|---|---|---|---|---|---|
| Two Monitors | Hydravison | Slide | 11.222 | .479 | 10.199 | 12.097 |
| | | Spreadsheet | 16.926 | .511 | 15.913 | 17.939 |
| | | Text | 13.778 | .430 | 12.925 | 14.631 |
| Three Monitors | Single | Slide | 10.481 | .483 | 9.525 | 11.438 |
| | | Spreadsheet | 15.815 | .528 | 14.769 | 16.861 |
| | | Text | 11.259 | .502 | 10.265 | 12.254 |
| | Multi-screen | Slide | 10.574 | .514 | 9.555 | 11.593 |
| | | Spreadsheet | 17.815 | .498 | 16.828 | 18.802 |
| | | Text | 13.222 | .497 | 12.236 | 14.208 |
| | Hydravision | Slide | 11.704 | .479 | 10.754 | 12.653 |
| | | Spreadsheet | 17.333 | .511 | 16.321 | 18.346 |
| | | Text | 13.796 | .430 | 12.943 | 14.650 |

**Table 14: Conditions by screens by tasks means, standard errors, and confidence intervals for Accuracy.**

**Table 15**

| | Task | Single Mean | Multi Mean | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 10.130 | 11.315 | 1.185 | 12 |
| | Spreadsheet | 16.315 | 16.630 | 0.315 | 2 |
| | Text | 10.500 | 12.796 | 2.296 | 22 |
| Three Monitors | Slide | 10.481 | 10.574 | 0.093 | 1 |
| | Spreadsheet | 15.815 | 17.815 | 2.000 | 13 |
| | Text | 11.259 | 13.222 | 1.963 | 17 |

**Table 15: Comparison of SS screen Accuracy means with MS Accuracy means, difference, and percent of change.**

**Table 16**

| | Task | Single Mean | Hydravision | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 10.130 | 11.222 | 1.092 | 11 |
| | Spreadsheet | 16.315 | 16.926 | 0.611 | 4 |
| | Text | 10.500 | 13.778 | 3.278 | 31 |
| Three Monitors | Slide | 10.481 | 11.704 | 1.223 | 12 |
| | Spreadsheet | 15.815 | 17.333 | 1.518 | 10 |
| | Text | 11.259 | 13.796 | 2.537 | 22 |

**Table 16: Comparison of SS screen Accuracy means with HV Accuracy means, difference, and percent of change.**

**Table 17**

| Configuration | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| Single | Slide | 10.306 | .341 | 9.629 | 10.982 |
| | Spreadsheet | 16.065 | .373 | 15.325 | 16.805 |
| | Text | 10.880 | .355 | 10.177 | 11.583 |
| Multi-screen | Slide | 10.944 | .363 | 10.224 | 11.665 |
| | Spreadsheet | 17.222 | .352 | 16.524 | 17.920 |
| | Text | 13.009 | .352 | 12.312 | 13.707 |
| Hydravison | Slide | 11.491 | .328 | 10.841 | 12.141 |
| | Spreadsheet | 17.130 | .361 | 16.413 | 17.846 |
| | Text | 13.787 | .304 | 13.184 | 14.390 |

**Table 17: Means, standard errors, and confidence intervals for SS, MS, and HV configurations by tasks over Accuracy.**

19 through 21provide the cell means comparisons that contribute to the data record. Table 22 presents the screen configuration means for each task in order to investigate the significant Screens by Task interaction.

The data in Table 22 shows a consistent advantage for multi-screen configurations across all tasks in terms of shorter average time per edit. These differences are significant for all but the SS to MS comparison for the slide task ($\alpha$ = .37). There are no significant differences between MS and HV means, although the pattern of HV being more effective in slide and text tasks is repeated. In terms of absolute values, multi-screen configurations (MS and HV combined) result in a savings of 2.2 seconds per slide edit, 3.2 seconds per spreadsheet edit and 6.7 seconds per text edit.

## Analysis and Results: Usability Data

### Analysis

Data from the usability questionnaires that were collected at the end of every task performance

**Table 18**

| Factors | F-Test | Degrees of Freedom | Significance |
|---|---|---|---|
| Screens by Tasks by Condition | 1.322 | 4, 424 | .261 |
| Screens by Condition | .701 | 2, 212 | .497 |
| Tasks by Condition | .488 | 2, 212 | .615 |
| Screens by Task | 5.742 | 4, 424 | .000 |
| Screens | 23.452 | 2, 212 | .000 |
| Tasks | 282.492 | 2, 212 | .000 |
| Conditions | .006 | 1, 106 | .940 |

**Table 18: Means, standard errors, and confidence intervals for SS, MS, and HV configurations by tasks over Time per Completed Edit.**

**Table 19**

| Monitors | Screens | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| Two Monitors | Single | Slide | 28.167 | 1.644 | 24.908 | 31.426 |
| | | Spreadsheet | 15.120 | .886 | 13.363 | 16.878 |
| | | Text | 28.739 | 1.523 | 25.719 | 31.759 |
| | Multi-screen | Slide | 25.299 | 1.742 | 21.846 | 28.752 |
| | | Spreadsheet | 12.525 | .680 | 11.177 | 13.873 |
| | | Text | 22.479 | 1.352 | 19.799 | 25.158 |
| | Hydravison | Slide | 25.223 | 1.533 | 22.183 | 28.262 |
| | | Spreadsheet | 12.911 | .777 | 11.370 | 14.453 |
| | | Text | 21.371 | .926 | 19.535 | 23.207 |
| Three Monitors | Single | Slide | 28.081 | 1.644 | 24.823 | 31.340 |
| | | Spreadsheet | 16.023 | .886 | 14.265 | 17.780 |
| | | Text | 28.360 | 1.523 | 25.340 | 31.380 |
| | Multi-screen | Slide | 28.667 | 1.742 | 25.214 | 32.119 |
| | | Spreadsheet | 11.913 | .680 | 10.564 | 13.261 |
| | | Text | 22.425 | 1.352 | 19.745 | 25.105 |
| | Hydravision | Slide | 24.223 | 1.533 | 21.183 | 27.262 |
| | | Spreadsheet | 12.044 | .777 | 10.502 | 13.585 |
| | | Text | 20.925 | .926 | 19.089 | 22.761 |

**Table 19: Conditions by screen configurations by tasks means, standard errors, and confidence intervals for Time per Completed Edit.**

**Table 20**

| | Task | Single Mean | Multi Mean | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 28.167 | 25.299 | 2.868 | 10 |
| | Spreadsheet | 15.120 | 12.525 | 2.595 | 17 |
| | Text | 28.739 | 22.479 | 6.26 | 22 |
| Three Monitors | Slide | 28.081 | 28.667 | -0.586 | -2 |
| | Spreadsheet | 16.023 | 11.913 | 4.11 | 26 |
| | Text | 28.360 | 22.425 | 5.935 | 21 |

**Table 20: Comparison of SS screen Time per Completed Edit means with MS Time per Completed Edit means, difference, and percent of change.**

**Table 21**

| | Task | Single Mean | Hydravision | Difference | Percent Change |
|---|---|---|---|---|---|
| Two Monitors | Slide | 28.167 | 25.223 | 2.944 | 10 |
| | Spreadsheet | 15.120 | 12.911 | 2.209 | 15 |
| | Text | 28.739 | 21.371 | 7.368 | 26 |

**Table 21 (Continued)**

| | | | | | |
|---|---|---|---|---|---|
| Three Monitors | Slide | 28.081 | 24.223 | 3.858 | 14 |
| | Spreadsheet | 16.023 | 12.044 | 3.979 | 25 |
| | Text | 28.360 | 20.925 | 7.435 | 26 |

**Table 21: Comparison of SS screen Time per Completed Edit means with HV Time per Completed Edit means, difference, and percent of change.**

**Table 22**

| Configuration | Tasks | Mean | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| Single | Slide | 28.124 | 1.162 | 25.820 | 30.428 |
| | Spreadsheet | 15.572 | .627 | 14.329 | 16.814 |
| | Text | 28.549 | 1.077 | 26.414 | 30.685 |
| Multi-screen | Slide | 26.983 | 1.232 | 24.541 | 29.425 |
| | Spreadsheet | 12.219 | .481 | 11.266 | 13.172 |
| | Text | 22.452 | .956 | 20.557 | 24.347 |
| Hydravison | Slide | 24.723 | 1.084 | 22.573 | 26.872 |
| | Spreadsheet | 12.478 | .550 | 11.388 | 13.567 |
| | Text | 21.148 | .655 | 19.850 | 22.446 |

**Table 22: Time per Completed Edit means, standard errors, and confidence intervals for each screen configuration by task.**

(9 questionnaires per respondent) were analyzed in a tasks by screens repeated measures design that examined differences across tasks and screens for each of effectiveness, comfort, ease of learning, productivity, mistake recovery, task tracking, task focus, and ease of movement across sources.

In order to determine if respondent perceptions of usability differed across screens and task types, a comparison of the three screen configurations and three task types was conducted separately for each of the eight items. Figure 2 presents the design as replicated across each item. This design allows the analysis of the relationship between screen configuration and task on each of the respondents' judgments of usability. Based on our initial suppositions, it was hypothesized that multi-screen configurations would score higher on each item than the single screen. Two monitor and three monitor multi-screen configurations were used to further strengthen the potential understanding of multi-screen effects. The comparison of multi-screens with and without screen management software was considered exploratory and no hypotheses were developed.

**Results**

As hypothesized, multi-screen configurations scored significantly higher in usability than the single screen on every measure in every task. HV means were generally not significantly different from MS means on all measures but varied in direction of difference across tasks. Table 23 presents the means for each item across each screen configuration.

Differences in items showed the effect of screen configurations. In single screen, task tracking was significantly lower than any other item and ease of learning was significantly higher than any other. In

**Figure 2**

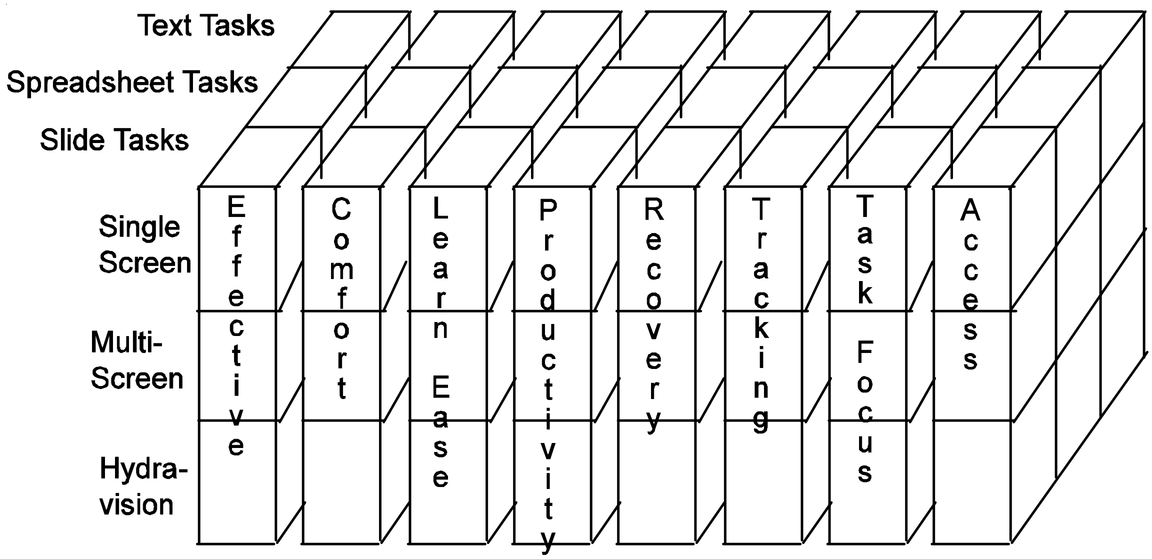


**Figure 2: Analysis of Variance design for screens by task type for each item.**

multi-screen, task tracking was also significantly lower than any other item, while accessibility was the highest, significantly higher than all but ease of learning. Hydravision means showed task tracking as significantly lower than all other items and accessibility as highest. Accessibility was significantly higher than mistake recovery, productivity, and comfort as well as task tracking.

Table 23 can also be used to calculate the changes in respondent judgments concerning screen configuration usability by using the single screen score and the average of the two multi-screen scores. In this analysis, multi-screens are seen as 29 percent more effective, 24 percent more comfortable, 17 percent easier to learn, 32 percent quicker to productivity, 19 percent easier for mistake recovery, 45 percent easier to track tasks, 28 percent better for task focus, and 38 percent easier for moving among sources.

## Discussion: Performance

This section first considers the central question of the effectiveness of multiple screens, briefly looks at the differences among tasks, the differences between conditions, then considers the interaction between screens and tasks, and finally examines the circumstances under which particular screen configurations should be adopted.

### Screens

The effect of screen configurations is quite clear. Respondents were able to get on task quicker, do the work faster, and get more of the work done with fewer errors in multi-screen configurations than with a single screen. The gains are solid: 6 percent quicker to task, 7 percent faster on task, 10 percent more production, 16 percent faster in production, 33 percent fewer errors, and 18 percent faster in errorless production. Equally impressive is that these gains were achieved by turning on an extra monitor or two and providing five minutes of training.

The value added by the screen management tool, Hydravision is subtle. It did not reach significance, but it was consistent and showed its greatest strength in controlling errors. Very little of the features of this software were used in this study, because of the nature of the tasks and measurements involved.

### Tasks

Without question, the most difficult task for most

**Table 23**

| Item | Screens | Means | Standard Error | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| Effective | Single | 6.419 | .174 | 6.074 | 6.763 |
| | Multi-screen | 8.309 | .147 | 8.018 | 8.600 |
| | Hydravision | 8.203 | .156 | 7.894 | 8.513 |
| Comfortable | Single | 6.549 | .177 | 6.198 | 6.900 |
| | Multi-screen | 8.216 | .149 | 7.921 | 8.511 |
| | Hydravision | 8.008 | .168 | 7.676 | 8.340 |
| Learning Ease | Single | 7.182 | .166 | 6.852 | 7.511 |
| | Multi-screen | 8.385 | .130 | 8.126 | 8.643 |
| | Hydravision | 8.388 | .139 | 8.112 | 8.663 |
| Productivity | Single | 6.230 | .182 | 5.869 | 6.591 |
| | Multi-screen | 8.247 | .143 | 7.964 | 8.530 |
| | Hydravision | 8.157 | .171 | 7.818 | 8.495 |
| Mistake Recovery | Single | 6.813 | .163 | 6.490 | 7.136 |
| | Multi-screen | 8.096 | .134 | 7.831 | 8.361 |
| | Hydravision | 8.081 | .152 | 7.779 | 8.382 |
| Task Tracking | Single | 5.311 | .218 | 4.878 | 5.743 |
| | Multi-screen | 7.724 | .202 | 7.323 | 8.124 |
| | Hydravision | 7.717 | .221 | 7.279 | 8.156 |
| Task Focus | Single | 6.432 | .195 | 6.046 | 6.818 |
| | Multi-screen | 8.211 | .147 | 7.920 | 8.502 |
| | Hydravision | 8.299 | .145 | 8.012 | 8.587 |
| Accessibility of Sources | Single | 6.193 | .203 | 5.791 | 6.594 |
| | Multi-screen | 8.612 | .129 | 8.357 | 8.868 |
| | Hydravision | 8.481 | .153 | 8.178 | 8.783 |

**Table 23: Means, standard errors, and confidence intervals for Usability items by screens.**

respondents was the slide editing task. Respondents reported the least experience with the application (mean of 1.25 on a scale of 0-3). These self-reports were confirmed in actual observations. A common failing was the inability to recognize substantial content differences among slides that had common backgrounds. Further, interview responses indicated that respondents were frustrated by the awkwardness of the application's editing protocols. When the slide task was removed from analysis, the efficiencies of time to task and time through task rose from 6 and 7 percent to 9 and 10 percent respectively.

The spreadsheet task for its part showed the shortest times to completion. Respondents reported slightly more experience with the spreadsheet application

(M = 1.40) than with the slide application. Observational notes show that respondents benefited from the spreadsheet application's ease of editing. When errors were made, they were generally entries in the wrong cell. Most commonly an entire row of entries were shifted up or down, accounting for the relatively large number of total errors in the individual edits.

The text task showed the fewest errors but also the lowest proportion of completed edits. Respondents indicated substantial experience with the application (M = 2.22), but few had experience with editing across screens (most work from paper corrections to a screen). Observations indicate that the visual task of locating place from one screen to another was the key difficulty.

**Conditions**

Conditions represent whether a respondent completed the study using two monitors or three monitors. Interestingly, the study was designed to "naturally" fit a three-monitor display, but the three-monitor condition consistently showed no advantage over the two-monitor condition. Anecdotally, multi-screen users consider the three-monitor display to be optimum, but it did not show here. Observations and comments from interviews suggest that the size of the monitor interacts with the optimal number of screens. Drawing on the comments of one respondent, a highly experienced graphics editor, the 18-inch monitors were too large for a three screen display as one could not keep the entire display within the field of vision. It may be very useful to advance this study with one that uses a three 15-inch monitor configuration.

**Screens by Task**

The lesson learned in the screen by task interaction is that there appears to be an optimum level of experience with a task that maximizes the immediate effect of the adoption of multi-screens. Too little as in the slide task, and the inexperience is an overburden on the multi-screen effect. Too much, as in the text task, and the productive methods of single screen editing prove a worthier competitor to reduce the size of the effect. Both of these conditions are functions of the testing protocol. Respondents given the regular experience of editing slide presentations would eliminate many of their difficulties, and respondents given the regular experience in multi-screen editing would return the competition to a level field.

The greatest proportion of our respondents (95 %) work only in single screen whether at home, at school, or in the office. As multi-screens were more effective than single screen across all tasks on measures of both time and production, it is clear that there is little learning curve in the adoption of multi-screen configurations. The short run benefits of converting to a multi-screen set up should be immediate and the long term gains substantial.

**Performance Considerations for Adoption of Multi-screens**

This study was designed to simulate office tasks that involve the application of multiple sources of information to a final product. It was, therefore, specifically designed to be responsive to the characteristics of multi-screen displays. The evidence it generated and the recommendations provided here presume similar circumstances—work that involves the integration of multiple sources. In those circumstances, the evidence speaks clearly and convincingly that multi-screen configurations are preferable and make good economic sense.

But not all work involves multiple sources of information. The question can be raised as to what proportion needs to be multi-sourced to justify the expense of adding that additional display port and monitor. The simplest way to answer that question is to extrapolate from the time per edit measure. The evidence suggests a 16 percent savings in time for the same level of production. Over a year's time, one would save $3,840 in labor costs at a $12 per hour clerical wage. Costs for upgrading computers vary by platform, region, and industry. Done at the authors' location with PCs, the upgrade would cost approximately $800 (adding a $75 PCI display card, a $600 LCD monitor, and the installation labor).

The break even point is approximately 21 percent of the work. If more than 21 percent of the work involves the use of multiple sources of information, upgrading to multiple screens is cost effective. The reader is also reminded that the break even point will be lower (less than 20%) with a less experienced (rather than diversely experienced) work force and even lower (less than 17 %) with a highly experienced work force.

## Discussion: Usability

Usability results showed the consistency of a mantra: Multi-screens either with or without management software are reported as significantly more usable than single screens on measures of effectiveness, comfort, learning ease, time to productivity, mistake recovery, task tracking, task focus, and ease of source movement. Slide tasks were considered the most difficult; spreadsheet tasks the easiest. Further, the least proficient respondents moved immediately to the level of the most proficient in their evaluations. They were not intimidated by the introduction of multi-screen displays. The open-ended interview data confirmed the positive response to multi-monitor displays. Those data showed overwhelmingly more positive comments for multi-screen and for Hydravision than for single screen and indicated that both multi-screen and Hydravision would be more useable and more likely associated with positive affect. Unlike many technological improvements, the adoption of multi-screen configurations should be a positive experience for the workforce and highly preferred over single screen arrangements. It not only increases productivity; the work is also judged as easier to do.

## Summary and Conclusions

This study compared single screen computer display configurations with multi-screen displays without screen management software and with multi-screen displays with screen management software—ATI's Hydravision. The comparisons were made using three types of ordinary office editing tasks in slide, spreadsheet, and text applications.

Multi-screens fared significantly better than single screen on time and number performance measures. Respondents got on task quicker, did the work faster, and got more of the work done with fewer errors in multi-screen configurations than with a single screen. They were 6 percent quicker to task, 7 percent faster on task, generated 10 percent more production, were 16 percent faster in production, had 33 percent fewer errors, and were 18 percent faster in errorless production. These gains are achieved by turning on a monitor and five minutes of training. Nonetheless, some care must be taken in extrapolating these gains over three 5-minute tasks to time saved and production increases achieved over a 40 hour work week. Such gains depend on the nature of the work and the amount of time spent on task and on multi-screen tasks. There is a utility in replicating this study using the tasks integrated into a continuous work period rather than as separate episodes as done here.

Respondents considered multi-screen configurations significantly more useful than single screen on every usability measure. Multi-screens were seen as 29 percent more effective for tasks, 24 percent more comfortable to use in tasks, 17 percent easier to learn, 32 percent faster to productive work, 19 percent easier for recovery from mistakes, 45 percent easier for task tracking, 28 percent easier in task focus, and 38 percent easier to move around sources of information. These increases were immediate post-test gains. As always, long-term gains may be different.

There were no significant differences between two-monitor and three-monitor multi-screen configurations. There was some evidence, however, of a relationship between optimum monitor size and the optimum number of monitors. The gains provided by an additional monitor in a three-monitor array can apparently be offset if the monitors are too large, forcing the user to physically track across the screens with head movement. It is recommended that a three-monitor array with 15 to 17-inch monitors be tested. Testing should also investigate portrait and landscape orientations within the array.

Given the overwhelming consistency of both the performance and usability measures, multiple monitor configurations are recommended for use in any situation where multiple screens of information are an ordinary part of the work. There will be measurable gains in productivity, and the work will be judged as easier to do. In addition, because the gains are strong, multiple monitors are also recommended as cost effective where multi-screen tasks represent as little as 15 percent of the work for the highly competent, 17 percent for entry level competence, and 21 percent for the general work force.

The contemporary status of computer displays is poised on a moment of convergence as operating systems can now handle multiple monitors with some routine; display boards with multiple ports are readily available and inexpensive; and LCD monitors with reduced foot print, space, and energy requirements, as well as cost are now becoming the standard. This study demonstrates that multiple monitor arrays should also be a standard of the workplace.

## Notes

This research was initiated by Don Lindsay, Chief Investigator, Lindsay Research & Consulting, Inc .and sponsored by ATI Technologies Inc. in collaboration with Neil Rennert, Marketing Research Manager, and Richard Mulcahy, Senior Product Marketing Manager, and by NEC/Mitsubishi in collaboration with Christopher Connery, Director of Marketing. James Anderson was the Principal Investigator.

1. Hydravision, ATI's screen management software, was used, hence the HV acronym.

## References

Barua, A., Kriebel, C. & Mukhopadhyay, T. (1991). *Information Technology and Business Value: An Analytic and Empirical Investigation*, Austin, TX: University of Texas Working Paper, (May).

Binder, C. (2001). Measurement: A few important ideas. *International Society for Performance Improvement. Performance Improvement Journal. 40*(3), 20-28.

Bohannon, W.K. (2003). Double vision. *Emedia Magazine, 16*(5), 22-28.

Brown, R.M., & Ruf, B. (1989). Applying software design principles to accounting software: A direct manipulation approach. *Journal of Information Systems, 4*(1), 41-55.

Brynjolfsson, E., & Yang, S. (1996). Information technology and productivity: A review of the literature. *Advances in Computers, 43*, 179-214.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Delfino, E. (1993). Windows and multitasking. *Database Magazine, 16*(6), 87-91.

Dyson, P. (2002). New technology sparks innovation in displays. *Seybold Report Analyzing Publishing Technologies, 2*(12), 11-14.

Gerlach, J.H., & Kuo, F. (1991). Understanding human-computer interaction for information systems design. *MIS Quarterly, 15*(4), 527-549.

Grudin, J. (2003). *Primary tasks and peripheral awareness: A field study of multiple monitor use.* Redman, WA: Microsoft Research.

Jorgenson, D.W., & Stiroh, K.J. (1999). Information technology and growth. *American Economic Review, 89*(2), 109-116.

Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7*(1), 57-58.

Lindsley, O.R. (1997). Performance is easy to monitor and hard to measure. In Kaufman, R., Thiagarajan, S., & MacGillis, P. (Eds.), *The Guidebook for Performance Improvement: Working with Individuals and Organizations* (pp. 519-559). San Francisco: Jossey-Bass/Pfeiffer.

Randall, N. (1999). Multiple monitors. *PC Magazine, 18*(16), 189.

St. John, M., Harris, W. C., & Osga, G. (1997). Designing for multi-tasking environments: Multiple monitors vs. multiple windows. *Proceedings of the Human Factors and Ergonomics Society* (pp. 1313-1317). Santa Monica, CA: The Society.

Sherry, L., & Wilson, B. (1996). Supporting human performance across disciplines: A converging of roles and tools. *Performance Improvement Quarterly, 9*(4), 19-36.

Stolovich H.D., & Keeps, E.J. (1992). What is human performance technology? In H.D. Stolovich and E.J. Keeps (Eds.) *Handbook of Human Performance Technology* (pp. 3-13). San Francisco: Jossey-Bass.

Vellotti, J.P. (2001). XP. *PC Magazine, 20*(9), 120.