

Visual Search and Reading Tasks Using ClearType and Regular Screen Displays: Two Experiments

Andrew Dillon, Lisa Kleinman, Gil Ok Choi, Randolph Bias
University of Texas at Austin
School of Information
1 University Station, D7000
Austin, TX 78712-1276
{adillon, kleinman, higilok, rbias}@ischool.utexas.edu

Summary

Two experiments comparing user performance on ClearType and Regular displays are reported. In the first, 26 participants scanned a series of spreadsheets for target information. Speed of performance was significantly faster with ClearType. In the second experiment, 25 users read two articles for meaning. Reading speed was significantly faster for ClearType. In both experiments no differences in accuracy of performance or visual fatigue scores were observed. The data also reveal substantial individual differences in performance suggest ClearType may not be universally beneficial to information workers.

1 Introduction

Improvements in image quality for text displays are assumed to offer meaningful benefits to users performing a variety of office tasks (Muter, 1996). Since most information tasks involve a mix of direct reading, searching, scanning, and manipulation of documents, it can prove difficult to tease out exactly where image quality will have the greatest effect. In an earlier study of user response to ClearType across editing, scanning and reading tasks, each involving progressively greater amounts of direct eye-on-text contact, Dillon et al. (2003) observed a 5% speed improvement for reading ClearType but no performance or preference differences for either the scanning or editing tasks.

An examination of user performance in that scanning study suggested that time spent typing the correct answer (in this case the text in a particular cell) was distorting the calculation of task speed, that is, the resulting data point measured both visual scanning and the physical activities of typing, components of the total activity for which image quality was irrelevant. Indeed not only was typing adding to the estimate of speed, but the requirement to specify the cell's text often necessitated several back and forth glances between target and keyboard. Similarly, for the editing task, much of the users' time was spent moving between screen and paper, and typing. In both these tasks, the impact of image quality was likely to be lessened, and its effects swamped by other factors.

The present experiments are intended to explore this issue further, particularly as it pertains to scanning, an activity that is highly visual. Specifically, we are interested in better controlling the user's interaction with the document so that only the visual act of interaction, rather than physical document manipulation or typing is involved. Unlike serial reading of text, scanning does not require users to follow a defined visual path, and the goal of target location involves different cognitive processes than text comprehension.

A further motivation for the present study was to replicate the reading performance effect observed previously. Participants in the previous experiment were noticeably fatigued at the end of the

experimental trial and it is possible that this effect impacted the reading speed scores collected during the last trial.

To explore these issues directly we designed two separate experiments. In the first, users scanned a set of spreadsheets for target information. The task was designed to require rapid scanning and time spent reading the task and answering did not contribute to the speed score, this overcoming one of the potential sources of confound observed in the earlier study. In the second experiment, users read a 2000 word article and answered a set of comprehension questions. Distinct user populations were employed for each experiment to control for any task or fatigue effects which may have impacted the earlier results.

Experiment 1: Scanning Task

2 Method

2.1 Participants

Participants were recruited through posted advertisements at the University of Texas at Austin and online job web sites. All participants met the following criteria:

- age 18 or older,
- English as a first language,
- self-reported familiarity using Internet Explorer
- self-reported as having 20/20, standard, or corrected vision
- self-reported as having no reading disabilities or color blindness

26 participants were scheduled for 1-hour time slots and paid \$10 for their involvement.

2.2 Experimental Design

The study employed a within-subjects (repeated measures) design with each participant completing ten unique scanning tasks in each of two conditions: ClearType and Regular text displays. After completing each set of tasks, the participant answered a visual fatigue questionnaire.

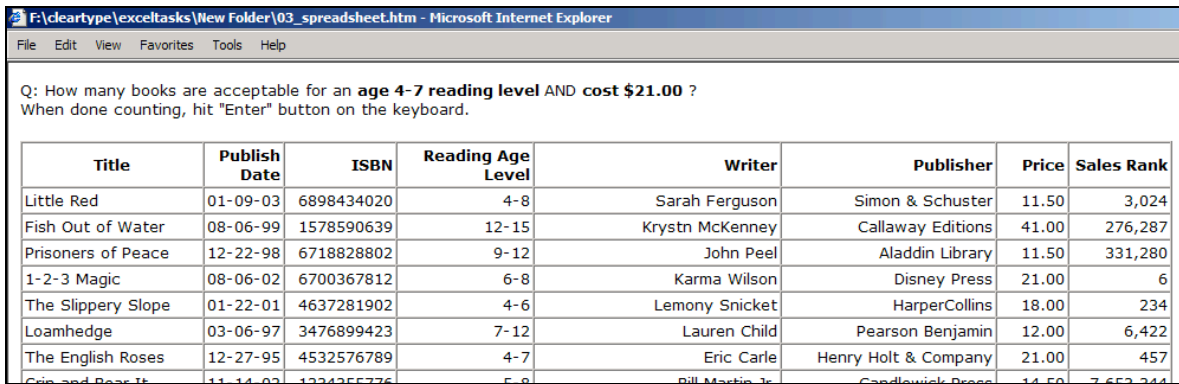
Performance was assessed through three dependent variables:

- 1) time to complete each scan
- 2) accuracy of answers
- 3) visual fatigue score

Each scanning task was presented using test software, controlled by the participant, which automatically captured time data and user responses. Accuracy data was based on the number of correct answers the participant provided; the total possible score was 20 points (one point per scan). The visual fatigue scores were based on Tyrrell's (1990) six scales of mental and physical fatigue with answers ranging on a seven point scale of strongly disagree to strongly agree. The wording of each of the visual fatigue questions was modified slightly to make sense for the scanning tasks (e.g. "working with computers" in the original Tyrrell questionnaire was changed to "answering these questions").

2.3 Experimental Material

For each task the subject was asked to determine the frequency of occurrence of target items in a spreadsheet display, see example in Figure 1.



Q: How many books are acceptable for an **age 4-7 reading level AND cost \$21.00** ?
When done counting, hit "Enter" button on the keyboard.

Title	Publish Date	ISBN	Reading Age Level	Writer	Publisher	Price	Sales Rank
Little Red	01-09-03	6898434020	4-8	Sarah Ferguson	Simon & Schuster	11.50	3,024
Fish Out of Water	08-06-99	1578590639	12-15	Krystn McKenney	Callaway Editions	41.00	276,287
Prisoners of Peace	12-22-98	6718828802	9-12	John Peel	Aladdin Library	11.50	331,280
1-2-3 Magic	08-06-02	6700367812	6-8	Karma Wilson	Disney Press	21.00	6
The Slippery Slope	01-22-01	4637281902	4-6	Lemony Snicket	HarperCollins	18.00	234
Loamhedge	03-06-97	3476899423	7-12	Lauren Child	Pearson Benjamin	12.00	6,422
The English Roses	12-27-95	4532576789	4-7	Eric Carle	Henry Holt & Company	21.00	457
Eric and the ...	11-14-03	1234567890	5-8	Bill Martin Jr.	Candlewick Press	14.50	7,653,244

Figure 1. Spreadsheet scan in Verdana text

All the text content in the scanning tasks came from actual web sites. Of the 20 test pages, 12 used Verdana, four were Courier, and four were Arial. No attempt was made to formally balance these fonts since all were deemed acceptable and typical by the experimental team but the tasks were matched for each condition so that the participant always completed a similar scan (same font, same length, similar content type) in both conditions. Order was counterbalanced for ClearType and Regular, as was the content within these conditions.

2.4 Equipment

Participants all used the same Dell Latitude C840 laptop, with the choice of using a mouse attached peripherally or the mouse-equivalents on the laptop keyboard. The screen size of the laptop was 15 inches, with a display setting of 1600x1200 pixels. ClearType was gamma tuned for the laptop to optimize its setting for this specific laptop. Internet Explorer 6.0 was used as the browser environment to present the test software and task.

2.5 Procedure

Each participant was seated in a closed room environment facing the laptop. The moderator explained the multi-part procedure for the session and then gave the participant practice scanning tasks. After completing two practice scans, participants proceeded through the first block of ten scans at their own pace.

For each task, the participant initiated the activity by reading the counting question presented on the test screen (e.g. "How many books were published in 1985?") and then clicking "enter" to make visible the spreadsheet screen to be searched and thereby starting the timer. As instructed, the participant again clicked "enter" when they had determined the answer, thus stopping the timer, at which point the test software then prompted the user to type their answer. After submitting the typed answer, the next task was presented on screen and the participant proceeded as before, until the first block of ten scans was complete. In this way, the timer only recorded the duration of the task for which the participant was scanning the spreadsheet.

At the end of this block, the stimuli and the moderator prompted the participants to answer the visual fatigue questionnaire (by hand on paper). The display condition was switched automatically by the test software while the participant filled in the fatigue scales and the second block of ten scans commenced as

before. Once more, at the end of this second block, the participant again completed the visual fatigue scales, after which the experiment ended.

3 Results

3.1 Summary of Results

Time, accuracy and visual fatigue scores were calculated for each participant and the summary data is presented in Table 1. Participants performed scans more quickly when ClearType was present, and there was noticeably greater variation in performance observed in the Regular condition. On average, participants were 24 seconds faster completing the scans with ClearType, a significant improvement of [$F_{(1,25)} = 5.07$, $p = .016$], see Table 2. There were no significant differences in terms of accuracy or visual fatigue.

	ClearType		Regular	
	Mean	Std. Dev.	Mean	Std. Dev.
Time (seconds)	279.54	43.46	303.81	79.42
Accuracy (max = 10 points)	8.46	1.07	8.73	1.22
Visual Fatigue (max = 7)	3.08	1.60	3.03	1.46

Table 1. Summary of scanning task data

Source	Sum of Squares	df	Mean Square	F	Sig.
DISPLAY	7656.94	1	7656.94	5.07	.016
Error (DISPLAY)	37744.56	25	1509.78		

Table 2. Repeated measures one-tailed ANOVA for scanning task

These results support the hypothesis that ClearType can improve performance in visually intensive scanning tasks. On average, where the eye-on-screen component of the task dominates, people performed 7.2% faster with ClearType than with Regular displays, with equivalent levels of accuracy and fatigue.

3.2 Mental and Physical Fatigue

The results of the perceived mental and physical fatigue scores for the scanning task are summarized in Table 3. For all scores, a greater value represents more perceived fatigue/strain (on a 7-point scale). No significant differences were observed between conditions.

Question	ClearType	Regular
Q1. My back and/or neck hurts from sitting in this position while answering the questions.	2.68	2.54
Q2. Answering these questions on a computer gives me a headache.	2.88	2.65
Q3. After answering these questions, my vision seems blurry when I look at distant objects.	3.84	3.69
Q4. I feel mentally fatigued right now.	2.56	2.54
Q5. My eyes feel strained right now.	4.12	4.00
Q6. Overall, answering these questions has made me feel fatigued.	2.88	2.73

Table 3. Mental and physical fatigue in the scanning task

3.3 Individual Differences

While the significant effect for display points to a robust benefit for ClearType on this type of visual scanning, the individual participant scores indicate that only 16 of the 26 participants performed better in the ClearType condition. However, for these people, the advantage of ClearType was pronounced, as shown in Table 4.

Difference Score Range	Number of Participants
0 to 30 seconds faster in ClearType	5
31 to 100 seconds faster in ClearType	7
Over 100 seconds faster in ClearType	4
0 to 30 seconds faster in Regular	5
31 to 100 seconds faster in Regular	5
Over 100 seconds faster in Regular	0

Table 4. Difference score groupings

Of the 16 who performed better with ClearType, four performed over 100 seconds faster, on a set of tasks that on average took around 300 seconds to perform. No equivalently large differences were observed for participants who were faster with Regular text. We can examine the full distribution of these individual differences graphically in Figure 2.

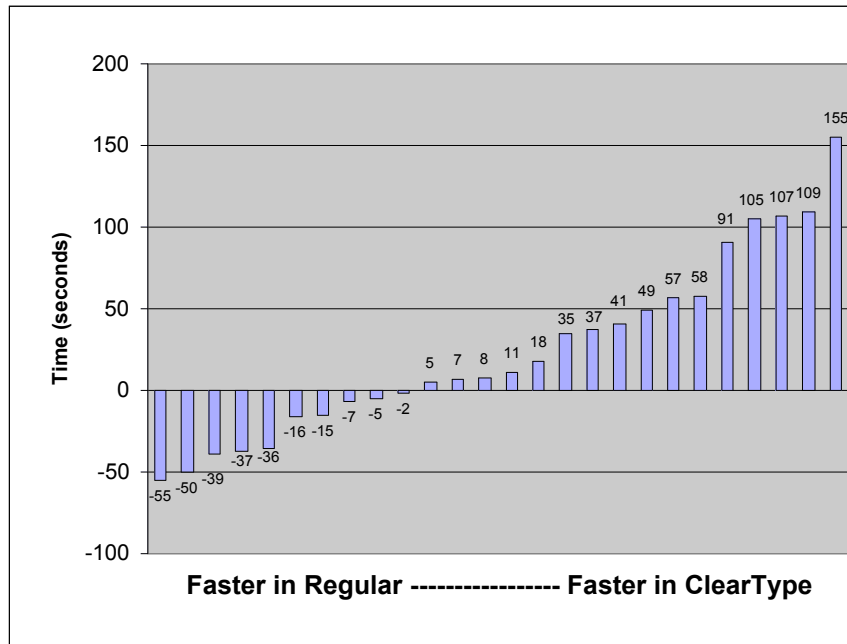


Figure 2. Difference scores in the scanning task

As this shows, once you remove those who demonstrated less than a 10% difference between conditions in either direction (30 seconds) then there are only 5 participants who were much better with Regular text but 12 who performed much better with ClearType. In fact, the mean performance benefit between displays was 53.5 seconds for the “ClearType-advantaged” users compared to a 26.2 seconds benefit for the “Regular-advantaged” (or “ClearType-disadvantaged”) users. These data seem to suggest that while ClearType can yield a general improvement in reading speed for many people, its benefits are not universal.

Experiment 2: Reading Task

4 Method

4.1 Participants

As before, participants were recruited through posted advertisements at the University of Texas at Austin.

All participants met the following criteria:

- age 18 or older,
- English as a first language,
- self-reported familiarity using Internet Explorer
- self-reported as having 20/20, standard, or corrected vision
- self-reported as having no reading disabilities or color blindness

25 participants were scheduled for 1-hour time slots and paid \$10 for their involvement.

4.2 Experimental Design

This experiment used a within-subjects 2 x 2 design for ClearType/Regular text and content of the reading material. Performance was assessed through four dependent variables:

- 1) time spent reading
- 2) accuracy of answers
- 3) visual fatigue score
- 4) preference score

Once more, test software was developed to enable automatic capture of time data for each task. To give us better insight into the process we divided the text into five screens of text, each approximately 400 words in length. To remove scrolling from the task, each screen was presented as a whole, and the participant pressed the space bar to move forward. The timer started running at the moment the participant displayed the first screen and also captured each screen's reading duration.

Accuracy data was based on the number of correct responses to questions on the material; the total possible score was 6 points. Five multiple-choice questions prompted for factual information from the article (as opposed to inferential reasoning questions); one point was given for each correct answer. Another additional point was calculated for a semi-structured question which asked the participant to list five themes or ideas that the article conveyed—1/5 of a point was given for each correct theme. The participant also responded to an open-ended question to summarize the article in three to four sentences, this was not calculated in the accuracy score but served as a check that the participant had actually read the text for comprehension.

Again, the visual fatigue scores were based on Tyrrell's (1990) six scales, modified as in the scanning task to suit this context. For this task we also added a preference score, calculated by presenting the participant with ClearType and Regular text side-by-side (counterbalanced across users). Participants rated the text sample which they preferred and how 'blurry' each image seemed on a 7-point scale, with 0 as the mid-point expressing no preference or choice, and each of the three points either side of this (from -3 to +3) representing degree of choice (slight, fair, or strong).

4.3 Experimental Material

Each participant read two articles of general interest: one about the history of newspapers, and the other about the etiquette of tipping. Reading difficulty and interest was deemed equivalent by the experimental team. Each article was approximately 2000 words long, and was divided over five screens and presented in Arial 12 pt font. Text and display orders were counterbalanced across participants. See Figure 3 for an example screen.

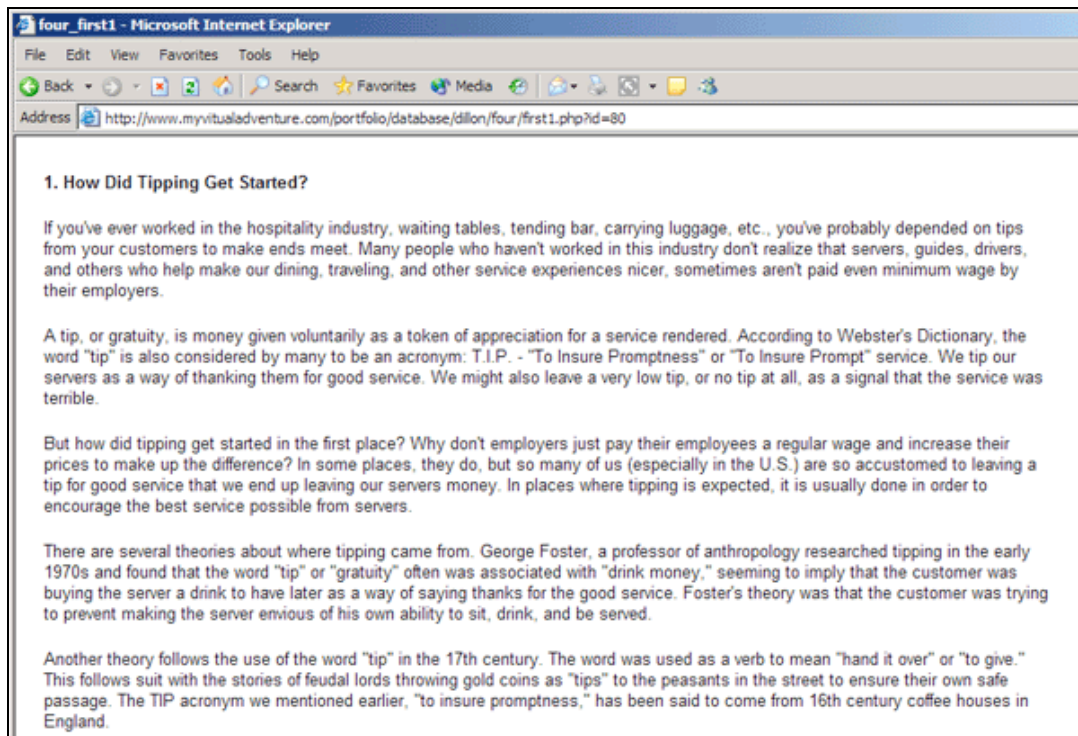


Figure 3. Reading task in ClearType Arial text

4.4 Equipment

Participants all used the same Dell Latitude C840 laptop, with the choice of using a mouse attached peripherally or the mouse-equivalents on the laptop keyboard. The screen size of the laptop was 15 inches, with a display setting of 1600x1200 pixels. ClearType was gamma tuned for the laptop to optimize its setting for this specific laptop. Internet Explorer 6.0 was used as the browser environment to present the test software and task.

4.5 Procedure

Each participant was seated in a closed room environment facing the laptop. The moderator explained the multi-part procedure for the session and demonstrated an example reading task to the participant who was allowed to ask questions and become comfortable with the procedure. Participants were asked to begin the first reading block by pressing the space bar and proceeding through the reading at their own pace.

After reading five screens (one complete article) the timing stopped and the participant answered the multiple and short-answer questions presented by the test software. After answering these questions, the moderator prompted the participant to answer the visual fatigue questionnaire by hand on paper while the display condition was switched. The participant then proceeded with the second reading article and questions, and answered the visual fatigue questionnaire a second time. The moderator finished the test session by showing a comparison of the ClearType and Regular texts on one screen, asking the participant to rate preference and perceived blurriness for each. On completion of these scales the experiment ended.

5 Results

5.1 Summary of Results

The summary data from all participants across all dependent variables is shown in Table 5. On average, participants read 31 seconds (5.65%) faster with ClearType. This is a significant result using a one-tailed repeated measures ANOVA [$F_{(1, 24)} = 3.12$, $p < .05$]. There were no significant effects for accuracy or visual fatigue.

	ClearType		Regular	
	Mean	Std. Dev.	Mean	Std. Dev.
Time (seconds)	548.60	184.00	579.32	190.37
Accuracy (max = 6 points)	5.06	.90	5.30	.70
Visual Fatigue (max = 7)	3.11	1.77	3.01	1.72

Table 5. Summary of reading task data

Source	Sum of Squares	df	Mean Square	F	Sig.
DISPLAY	11796.48	1	11796.48	3.12	.045
Error (DISPLAY)	90691.52	24	3778.81		

Table 6. Repeated measures one-tailed ANOVA for reading task

5.2 Reading Speed Across Text Blocks

Figure 4 shows the comparison of task completion time between ClearType and Regular text across each of the five screens (blocks) of text. Participants consistently performed faster with ClearType across all five screens, indicating a steady performance benefit for most users.

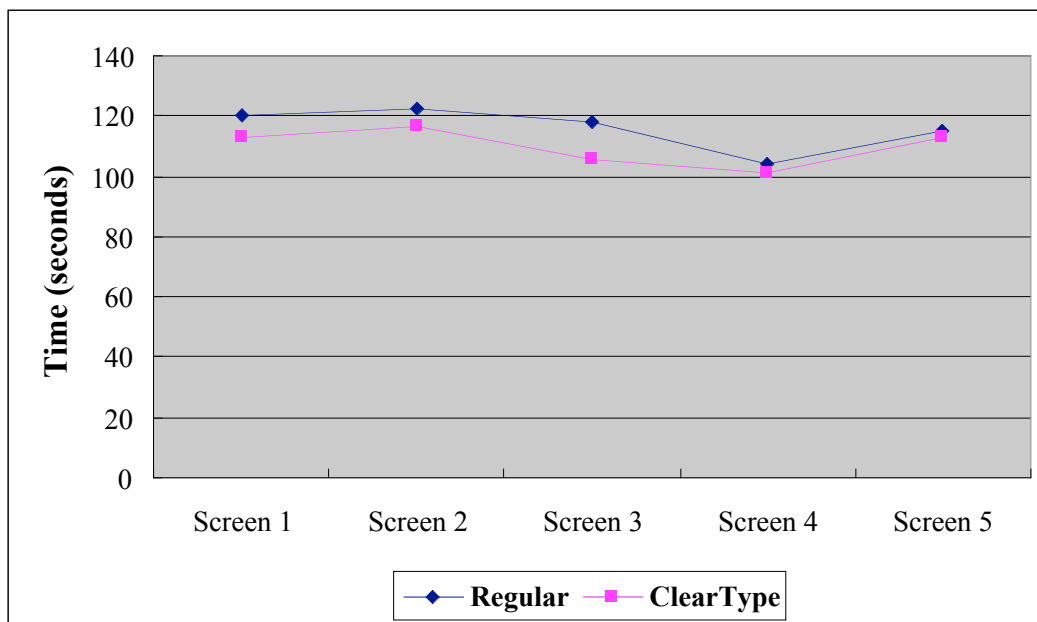


Figure 4. Reading times for each screen of text

The results of mental effort and visual fatigue score are summarized in Table 7. For all scores, a higher score represents greater perceived and physical demand or effort. Once again, no significant differences were observed between display conditions.

Question	ClearType	Regular
Q1. My back and/or neck hurts from sitting in this position while answering the questions.	2.69	2.58
Q2. Answering these questions on a computer gives me a headache.	2.88	3.15
Q3. After answering these questions, my vision seems blurry when I look at distant objects.	3.31	3.15
Q4. I feel mentally fatigued right now.	2.77	2.69
Q5. My eyes feel strained right now.	4.15	3.62
Q6. Overall, answering these questions has made me feel fatigued.	2.85	2.88

Table 7. Mental and physical fatigue in the reading task

5.3 Preference and Blurriness Scores

Overall, more participants strongly preferred ClearType and rated it less blurry when comparing it to regular text. Tables 8 and 9 show the number of participants for each of the scales grouped into categories of preference and blurriness. There is a significant negative correlation between preference and blurriness scores ($r = -.79, p < .01$), as expected.

Preference	ClearType	Regular
Slight	1	2
Fair	3	6
Strong	9	1
Total	13 (55%)	9 (45%)

Table 8. Preference groupings – Number of participants

Blurriness	ClearType	Regular
Slight	2	3
Fair	7	6
Strong	0	4
Total	9 (45%)	13 (55%)

Table 9. Blurriness groupings – Number of participants

5.4 Individual Differences

If we investigate individual differences in reading speed, as before, we find again that a proportion of participants were faster in the Regular condition. Table 10 summarizes these difference scores in global terms.

Difference Score Range	Number of Participants
0 to 50 seconds faster in ClearType	2
51 to 100 seconds faster in ClearType	10
Over 100 seconds faster in ClearType	4
0 to 50 seconds faster in Regular	4
51 to 100 seconds faster in Regular	3
Over 100 seconds faster in Regular	2

Table 10. Difference score groupings in the reading task

Again, viewing these graphically (see Figure 5) we can see that the differences are more pronounced in one direction. Mean improvement scores for the ‘ClearType-advantaged’ users were 85.75 seconds compared to a 65.77 seconds decrement for the ‘ClearType-disadvantaged’ users.

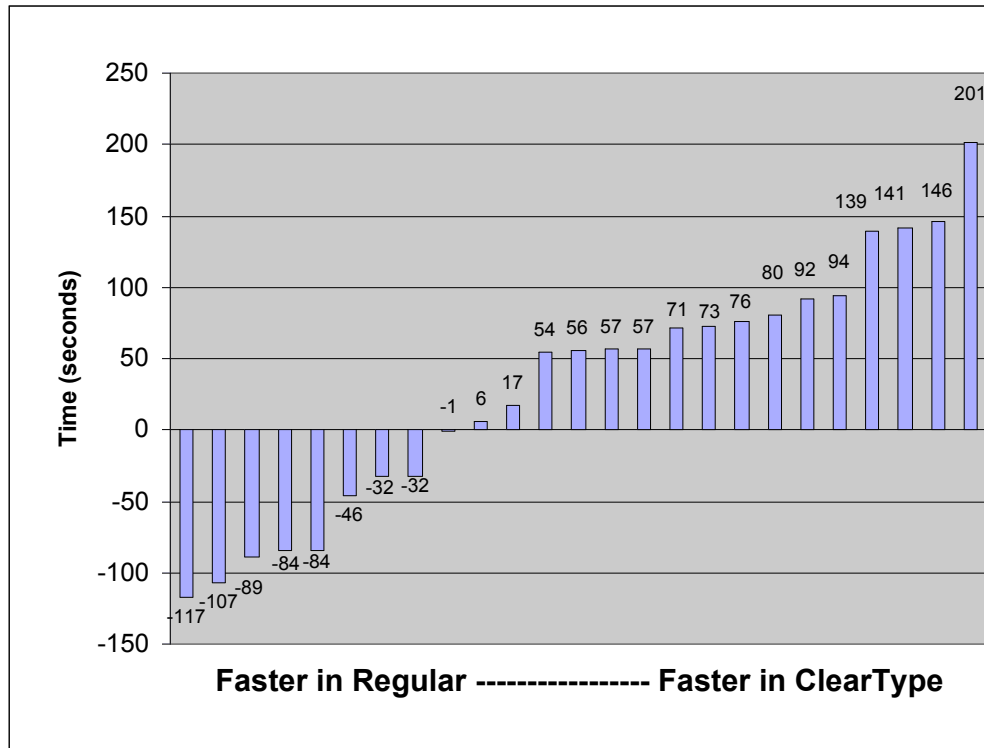


Figure 5. Difference scores in the reading task

These data again suggest that where ClearType helps, it does so in a more pronounced fashion than it hinders. However, once gain, the benefits are not universal.

6 Conclusions

ClearType-enhanced text produces a significant improvement in performance times for tasks involving extended eye-on-text interactions for most users. For both reading and visual scanning tasks of the type employed here, which were designed to simulate routine office tasks for information workers, the recommendation to use ClearType is supported. Across 51 users in this study, 32 witnessed improvements in their speed compared to regular text display. Furthermore, the size of this performance improvement seen by these users was more pronounced than the decrement witnessed by 19 users who were faster on regular displays.

The data from the reading task alone confirm the earlier finding from Dillon et al. (2003) which showed a 5.1% speed advantage to ClearType over Regular displays. The present study examined reading alone, not as a part of a combined task block, and it showed a 5.65% advantage to ClearType. In combination (covering more than 60 users), these data offer substantial support that reading speed is improved by the image quality enhancements of ClearType.

The performance on the visual scanning task was also improved by use of ClearType. While the original 2003 study had not shown an effect, the manipulations in the present study enabled us to gain a better isolated measure of visual scanning without other task components intruding, and here we see speed was improved an average of 7.2% by ClearType, as originally hypothesized.

Both these tests support the argument that image quality improvements have maximum benefit for tasks where the user is spending large proportions of the tasks with their eyes on the text. Where tasks are structured around other activities, moving between screen and paper, manipulating documents or complex typing, the effects of image quality improvements are likely swamped by other factors. Since information work consists of many interrelated physical, perceptual and cognitive activities, the impact of image quality improvements will not always be pronounced.

The issue of individual differences remains a source of interest. With some 19 users experiencing some disadvantage in the ClearType condition there is a need to determine what factors may be at work here. Since we have no demographic or perceptual processing data here by which to compare participants we can only speculate on potential acuity or related visual processing factors that may be at work. However, to check for the possibility of the type of order effect we reported in our previous study, we conducted a post-hoc analysis with display order as a variable. Interestingly, even though order was fully counter-balanced, there was an interaction effect for display condition with order for the scanning task ($F_{(1, 22)} = 10.95, p < .01$) and a three-way interaction effect between order, display and content for the reading task ($F_{(1, 22)} = 9.20, p < .01$).

In both tasks, there is no evidence of a practice effect (participants were not faster in the second task block than the first) but all four of the participants who were more than 100 seconds faster with ClearType in the scanning task happened to be exposed to ClearType before Regular Text. It is impossible to determine accurately with these data if this is significant but the effect may point to different impacts resulting from the direction of the transition to or from ClearType. In other words, it is possible that ClearType produces different effects in users when it is turned off, and they are required to read Regular text than when it is turned on after they have spent time reading Regular text. This is an issue for further study.

References

Dillon, A., Kleinman, L, Bias, R. & Choi, G. (2003) *Editing, searching, and reading electronic text with ClearType and regular displays: A comparative analysis*. Unpublished manuscript.

Muter, P. (1996) Interface design and the optimization of reading of continuous text. In H. van Oostendorp and S. de Mul (eds) *Cognitive Aspects of Electronic Text Processing*, Norwood NJ: Ablex, 161-180.

Tyrrell, R. & Leibowitz, H. (1990). The relation of vergence effort to report of visual fatigue following prolonged near work. *Human Factors*, 32(3), 341-357.