

Augmented Reading: An Experimental Study on Presenting Additional Information Without Penalty

Eric Bahna¹

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052, USA
ebahna@microsoft.com

Robert J.K. Jacob

Department of Computer Science
Tufts University
Medford, MA 02155, USA
jacob@cs.tufts.edu

ABSTRACT

We present a new interaction technique for computer-based reading tasks. Our technique leverages users' peripheral vision as a channel for information transfer by using a video projector along with a computer monitor. In our experiment, users of our system acquired significantly more information than did users in the control group. The results indicate that our technique conveys extra information to users nearly "for free," without adversely affecting their comprehension or reading times.

Author Keywords

Peripheral vision, eye tracking, reading, interaction technique evaluation, human-computer interaction, lightweight interaction, non-command interfaces, geographical information

ACM Classification Keywords

Interaction styles, evaluation/methodology

INTRODUCTION

We demonstrate a novel, interactive display technique that leverages peripheral vision to communicate additional information to a user engaged in a reading task. We wish to convey the additional information through peripheral images without distracting the user from his or her primary focus of reading the text. We base our design on lightweight, or non-command, user interaction [3, 7] because this approach allows us to present information to the user in a background or peripheral mode. Our design is motivated by a vision of nearly "free" ancillary information.

Consider a digital library about ancient Greece that includes a sentence such as, "Ictinus was the architect of the

Parthenon." Environmental, geographical, or spatial information about places in the text, such as the Parthenon, would enhance the reader's understanding of such a text, particularly a modern reader who is less familiar with the Parthenon than the author's contemporaries would have been. A conventional solution might provide a hypertext link in the text ("Ictinus was the architect of the Parthenon [click here for picture of Parthenon]") On a typical website, selecting the link would display the picture in the main browser window, replacing the text the user was reading; a "Back" button or other link would then take the user back to the text to continue reading. This method interrupts the reader's primary task, and is thus precisely the wrong approach. Our design suits situations in which the text is primary, and the environmental information is a relatively minor, ancillary piece of knowledge. We aim to keep the reader on his or her primary task (reading) while we simply provide him or her with some small amount of environmental information in the background. We do this with minimal distraction from the reading, without requiring any explicit action on the user's part, and without using up screen space that would normally be devoted to the text.

DESIGN

The motivation for our design comes from this scenario: Suppose a reader was enjoying a book while riding on a bus, which was driving through the same places mentioned in the text at the same time that they were mentioned. Even without looking up from the book, the reader would be able to get a rough sense of where he or she was at each point in the text – indoors, outdoors, urban, rural, daytime, nighttime – via peripheral vision through the bus windows.

¹ The work reported in this paper was carried out during the first author's affiliation with the Department of Computer Science at Tufts University.

Similarly, our interaction technique aims to convey general awareness of the environment of the text.

In our design, the user reads on a conventional display monitor, while the peripheral information is shown as a very large image on a wall just behind it. The peripheral image changes as the reader scrolls through the text in our system. We minimize distraction to the reader by gradually dissolving images from one to another.

Although the reader can look up at the wall at any time, our intention is that he or she need not ever look at it, but can simply receive information through peripheral vision, just like the imaginary reader riding in the bus. The reader thus uses his or her foveal vision only for the reading task. Peripheral vision is little used in reading, so we view it as an extra available channel in this setting and we exploit it for communicating an approximate sense of place. We believe that this combination of foveal and peripheral vision makes a fuller use of the reader's available attentional resources, thereby conveying additional information. We tested this hypothesis in our experiment below.

Implementation

Our peripheral border display system runs on a Windows PC with a graphics card that can support two displays. It runs as a Java application, with two windows, one driving the main LCD display monitor and one driving a video projector. Figure 1 shows a photograph of the two displays, with the entire main LCD screen (circled in the lower-center part of the photograph) used for reading text and the appropriate image related to the current text projected onto the surrounding wall. We use a 15" LCD monitor and a video projector image that is 62.5" high and 96" wide.²



Figure 1: Peripheral Border Display System: The LCD monitor is shown as the light rectangle in the middle of the lower half of the picture. The Polhemus transmitter is shown to the right of the monitor.

² This was the largest projected image that could be shown in our lab.

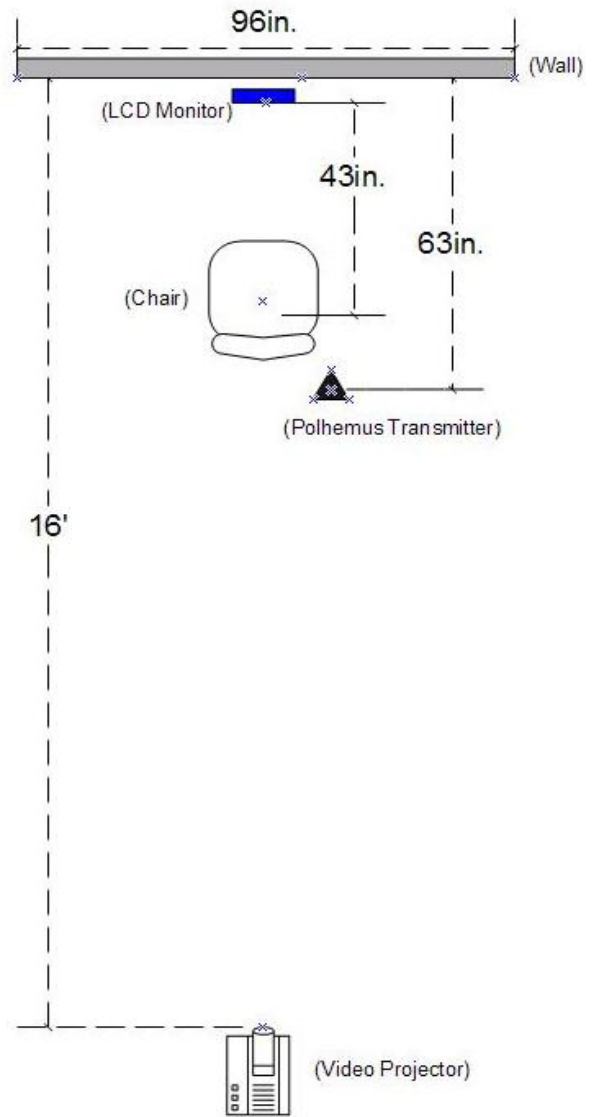


Figure 2: Diagram of Peripheral Border Display System (Overhead View)

The projector image is front-projected from behind the user's head onto a wall just behind the monitor using an inexpensive, 300 lumens, video projector. The projected image has a black cutout to cover the portion of it that would otherwise project over the LCD monitor. This prevents the projection from causing a glare on the monitor. A diagram of the system's setup is shown in.

The eye tracker system that we used was comprised of an ISCAN RK-726PCI Pupil/Corneal Reflection Tracking System along with an ISCAN RK-520PC Autocalibration System. All subjects wore a headband on which was mounted a lightweight eye tracking unit and a Polhemus 3Space FASTRAK magnetic tracker. Figure 3 shows the headband on a subject. The eye tracking unit consisted of a tiny camera, an infrared (IR) light source, and a dichroic mirror. The camera captured an image of the subject's left

eye and sent it to a second computer, separate from the computer that displayed the content. The second computer ran the ISCAN Headhunter Line of Sight Computation and Plane Intersection Software (version 1.51), which used the IR light source to determine the locations of the pupil and the corneal reflection. From those two pieces of information, it calculated the eye's line of sight. The ISCAN software also reads the head position and orientation, relative to the Polhemus transmitter that we placed in a fixed location in the lab (as shown in Figure 2). Combining data from the eye tracker and magnetic tracker, the ISCAN software determined the vector along which the subject is looking relative to a fixed point in the lab. Using the physical locations of the LCD monitor and the projected image as input, the software determined whether a particular eye gaze was directed at the LCD monitor, the projected image, or at some content-free location. We used this data to assess the extent to which the peripheral images attracted the foveal vision of subjects.

Our Java application displays the text in a window with a normal white background on the main computer screen. The background images are shown in a separate "window" that is displayed on the wall by the projector. To synchronize the text with the images, we marked up the text to add a new HTML-like tag at each point where a new image should be displayed. (We assume that the user will generally be reading near the vertical center of the window, not at the extreme top or bottom, and we synchronize the peripheral picture to the text that appears at that central area.) The images change by dissolving from one picture to the next.

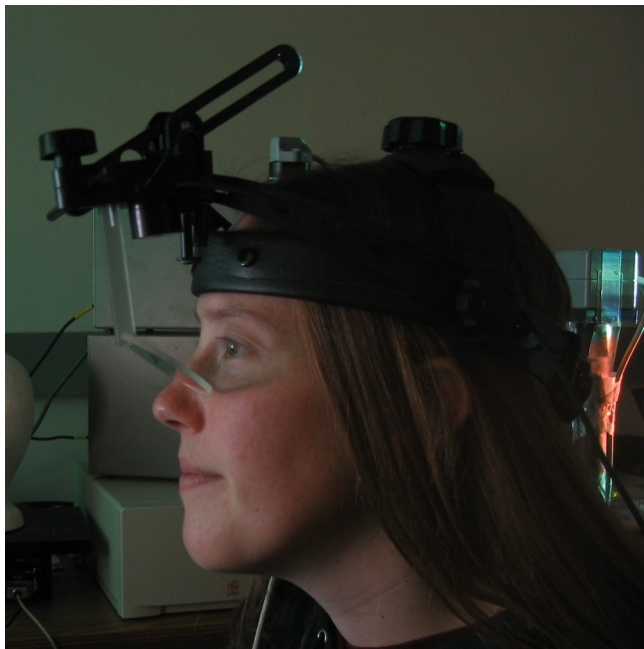


Figure 3: A Subject Wearing the Headband with the Eye Tracker and Polhemus Magnetic Tracker.

The peripheral image changed automatically after the subject scrolled through one window of text. In our configuration, the LCD monitor had a resolution of 1024x768 pixels, and the text window was 800x600 pixels with double-spaced text in a 15-point font. We chose a font size of 15 points rather than more conventional 10 or 12 points because subjects were reading from 40 to 48 inches away from the display. We had them read from such a distance so that they would have a good portion of the projection in their peripheral field of view.

Related Designs

We also designed and built two related interaction techniques that present additional information using a single monitor without the projector. One technique shows the extra information on the monitor in a border around the text. The other technique leverages the background behind the text. More information about these two can be found at our webpage for new human-computer interaction techniques for digital libraries [13]. In the black-and-white reproductions of screenshots provided below, the subtleties of the images may be difficult to detect. We strongly encourage readers to try the live prototypes on the webpage, which can provide a much truer impression of the user experience.

Our border display technique is based on the aforementioned analogy of reading on the bus. Rather than projecting an image on the wall around the monitor, here we display the corresponding image around the text window on the main monitor. To discourage users from focusing on the image, it is blurred to the point of capturing only a little more information than the overall color of the scene. As with the peripheral border display technique described above, the border image dissolves from one to the next as the user scrolls. A screenshot of the border display technique is shown in Figure 4.

For our background display technique, the analogy is that the user is reading text printed on a piece of clear plastic and can see, behind the plastic, a picture of the locale for that text. We thus display a window of text with the image as the background. As the story is read, the background changes to reflect the places mentioned at each point in the text. The image is toned down so as not to detract from the text, somewhat like a poster with text printed over faint background graphics. The image is also greatly blurred, even more so than in the border display technique. A screenshot of the background display technique is shown in Figure 5.

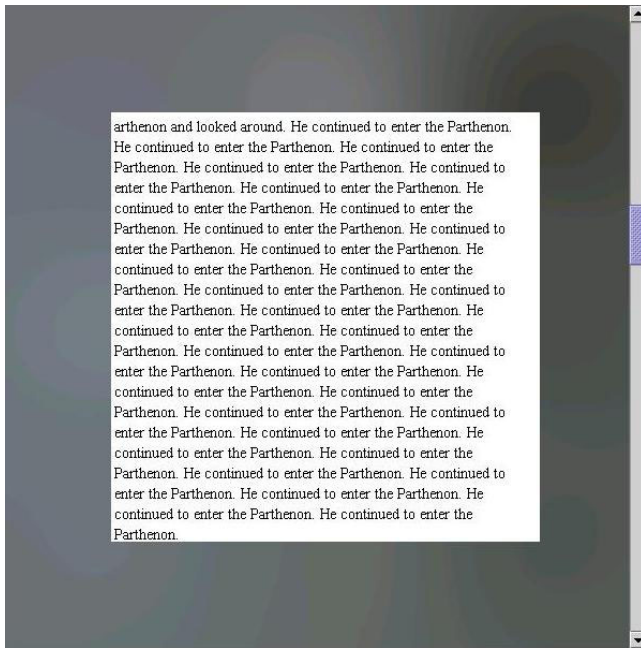


Figure 4: Screenshot of the Border Display Technique



Figure 5: Screenshot of the Background Display Technique

EXPERIMENT

We claim that our peripheral border display system can provide a user with extra information related to the text while he or she is reading, without having a negative impact on the reading task. To investigate this, we conducted an experiment with our peripheral border display. First, we measured the extent to which the system conveyed additional information to the user. Second, we evaluated potential drawbacks, such as an adverse effect on reading time or comprehension of the content. Third, we determined

more precisely how the additional information had been conveyed to the user.

Experimental Design

Our experiment utilized a completely randomized (between-subject) design with two conditions. Subjects in the experimental condition were asked to read content using the peripheral border display system as described above, with the text content presented on the LCD monitor and the images related to each part of the story projected on the wall around the monitor. Subjects in the control condition were shown the text content on the same LCD monitor, but the projector showed only a plain gray background rather than the images.

Subjects

Eighteen subjects (13 males and 5 females) volunteered for the study. All participants were college students from the Boston, Massachusetts, area, between the ages of 21 and 26, and were not computer novices. There was no compensation for participation.

Experimental Stimuli

Selection of appropriate content was critical to the success of the experiment. We designed the peripheral border display system for content that is primarily text and that also includes some pictures or graphics that convey ancillary information. We chose the following criteria for content selection:

1. The content is primarily text-based.
2. The content includes a sufficient number of pictures or graphics that contribute to the understanding of the text, but are not absolutely critical as they might be in an art history book, for example.
3. The text alone should take between 5 and 10 minutes to read.
4. It is not likely that any of our subjects will have read this content prior to our experiment.

The third criterion warrants some explanation. In order to determine whether the peripheral border display system affects reading time, we needed to allow our subjects' natural variability in reading times to manifest itself. If the text could be read too quickly, then we reasoned that there would be little variability in the reading times, thereby masking any subtle effect caused by the system. We also wanted to avoid an unusually long or burdensome task when recruiting subjects, so we wanted to keep the reading task from taking too long. We chose a target reading time of approximately five to ten minutes in order to balance both of these concerns.

We looked for content that already included text and images in one package, rather than identifying text and image content separately. Ultimately, we chose to use the contents of an illustrated children's book. We searched

through dozens of illustrated children's books and found that very few had both ample text to satisfy our time criterion and plenty of pictures. We selected the book [6] that appeared to strike the best balance of lengthy text and a sufficient number of pictures.

Our system is designed to change the peripheral images no more frequently than once per window of text. Our text content consumed between 7 and 8 windows, and for each window of text, we selected a pertinent image (7 images total). We tagged the text such that the picture would appear as the corresponding text was in the top half of the text window.

Procedure

Before the subject arrived, the projector was turned on and displayed a blank blue background on the wall. We learned from our pilot study that if the projector was turned off when subjects arrived then they became more interested in the projected image when it finally appeared. This then drew their attention away from the text on the LCD monitor, which is precisely what we wanted to discourage.

The subject was seated in a chair that was in a fixed location in the lab. The experimenter calibrated the eye tracker for the subject. Following calibration, the experimenter turned off the lights in the lab to make the projected images more visible. The experimenter explained that the subject's task was to read a story on the LCD monitor and answer questions about it afterwards. In order to give all subjects as similar a vantage point as possible, particularly with regard to the amount of the projection that was visible in their periphery, the experimenter instructed the subject to sit back in the chair. The experimenter informed subjects that they could safely ignore the projected material and that they would be timed while reading. The experimenter then launched the Java application that displayed the appropriate version of the story (experimental with images or control with fixed gray background), began recording eye data, and started a stopwatch.

In our pilot study, the instruction regarding the projected images was more stringent. Our pilot subjects were told explicitly to look only at the LCD monitor and not to look anywhere else in the room. Compliance with that strict instruction, combined with the fact that they knew their eyes were being tracked, appeared to require so much effort from our pilot subjects³ that their reading performance was adversely affected. This was particularly evident in the experimental condition, probably because the projected images were more tempting to look at than the gray background in the control condition. We therefore softened the wording of the instruction in an effort to keep the

³ When asked how easy or difficult it was to not look, pilot subjects made such comments as "it was challenging to not look" and "it was hard not to look."

subject more relaxed and reduce the artificial impact of the experiment on their reading performance. To mitigate the concern that this softer instruction might cause the subject to spend a significant amount of time looking at the images (because it was not expressly forbidden to look at them), we added a reminder that the subject's reading time was being monitored.

When the subject verbally indicated that the reading task was completed, the experimenter stopped recording the eye data, stopped the stopwatch, and closed the Java application. After turning the lights back on, the experimenter helped the subject remove the eye tracker headband. The subject was given the written memory test and then verbally asked some questions about his or her experience.

Data Collected

For each subject, we collected data from four different sources: the memory test, the subject's observed reading time, the eye tracker, and the verbal questions about the subject's experience.

Memory Test

Users were given a written test about the content after they were finished reading. Our goal was to devise a difficult test so that the scores would be as widely distributed as possible, thereby allowing us to better detect differences in performance. Since our text was originally written for a much younger audience than our subject pool, we made the test as lengthy and detailed as possible.

The memory test contained a total of 49 questions, which can be classified into 3 types:

1. Text questions (26 of 49): Questions whose answers were present only in the text.
2. Reinforced questions (10 of 49): Questions whose answers were present in both the text and the images.
3. Picture questions (13 of 49): Questions whose answers were present only in the images.

We were interested in the peripheral border display's effect on each of these types of questions. Of course, we believed that the experimental group would correctly answer more picture questions than the control group who never saw the pictures. However, we were concerned that the experimental group might be distracted by peripheral images, and perhaps retain less of the text content than the control group. We could detect that by comparing the two groups' performance on the text questions.

We decided to make some questions fill-in, or free response, while other questions were made multiple choice. Then, in order to take advantage of the fact that recall is more demanding than recognition [14], we sought to maximize the number of fill-in questions and minimize the number of multiple choice questions.

Multiple choice answers were provided for questions that met either of the following two criteria. First, the nature of the question was such that the universe of possible responses was very small (between 2 and 4 possible responses). For example: *When the characters were talking at the beginning of the story, were they inside or outside?* Second, the question could be answered at several different granularities and there was no clear, simple way to specify the desired granularity except by providing answer choices. For example: *During what time of day were the characters talking at the beginning of the story?* If this was a free response question, four subjects might all know the answer and yet their responses could be different: “6pm,” “dusk,” “afternoon,” and “evening.” We avoid this situation by instead presenting response choices “dawn,” “daytime,” “dusk,” and “nighttime.”

The distribution of questions according to their group and answer type is shown in Table 1. The majority of free-response questions (24 of 27) are text questions and the majority of multiple-choice questions are exclusively or partially based on the pictures (20 of 22). This is not surprising since we expect that the information communicated through peripheral vision will have less detail than would be afforded through foveal vision.

The distribution of picture questions might have a negative effect on the amount of information that the images appear to convey to subjects. As shown in Table 1, the overwhelming majority of picture questions (12 of 13) are multiple-choice. The difference between the experimental and control groups’ performance on the picture questions offered a metric of the extent to which the images conveyed information to the user. Because the vast majority of picture questions are multiple-choice and recognition is easier than recall, subjects in the control condition might correctly answer several picture questions despite not ever seeing the actual images. It is possible that this might reduce the difference in picture-question performance between the experimental and control conditions, thereby appearing to reduce the amount of information that the images conveyed.

Type of Question	Type of Answer		
	Free-response	Multiple-choice	ALL TYPES
Text	24	2	26
Reinforced	2	8	10
Picture	1	12	13
ALL TYPES	27	22	49

Table 1: Distribution of memory test questions by the type of question and the type of answer.

We attempted to mitigate this concern by offering an “unsure” response choice to all questions. Subjects were instructed to respond “unsure” to questions to which they were not certain of the answer. The “unsure” response also helped us mitigate the concern that the relative simplicity of the story allowed some answers to be logically deduced.

Reading Time

We measured how much time each subject spent reading the content. In the event that subjects in the experimental condition received more information than those in the control group (determined by their performance on the picture questions), we wanted to assess whether that additional information came with a cost of additional reading time. We also asked subjects for their estimates of how much time they spent reading.

Experiential Statements

Subjects were verbally asked for their level of agreement with statements about their experience after they read the content and completed the memory test. They responded using a 7-point Likert scale (1: strongly disagree ... 7: strongly agree) to statements about their enjoyment of the system, their enjoyment of the story, their level of fatigue, the extent to which they found the projected image to be helpful, and the extent to which they found it distracting.

Eye Tracker

We measured subjects’ eye movements during the experiment, to investigate the extent to which they actually looked at the “peripheral” display. From previous eye tracker studies, we expected that some subjects could not be properly calibrated with the eye tracker [10]. We collected the eye data to provide a richer set of data for that subset of subjects who could be successfully calibrated. Since we regarded the eye data as extra, we had subjects perform the experiment even if the eye tracker could not be calibrated with them. We were able to obtain eye tracking data for 14 of our 18 subjects (7 experimental and 7 control).

RESULTS

As expected, we found that subjects in the experimental group correctly answered more picture questions ($M = 6.67$, $SD = 2.35$) than did subjects in the control group ($M = 3.78$, $SD = 1.92$). We employed a single-factor, two-level, independent-measures analysis of variance (ANOVA), which suggested that the difference in performance was statistically significant: $F(1,16) = 8.169$, $p = .011$. We followed-up the ANOVA with the median test, a non-parametric statistic for significance. It too suggested that the two distributions were significantly different ($p = .024$) and did not share a common median. There was not a significant difference in performance on the text questions ($F[1, 16] = 0.187$, $p = .671$) or the reinforced questions ($F[1, 16] = 0.000$, $p = 1.000$). Performance on the memory test is summarized in Table 3.

More interestingly, we examined reading times between the two groups and found no significant difference between subjects seeing the peripheral projected image and those simply seeing the text: $F(1, 16) = 1.856, p = .192$. The observed mean and standard deviation was 364.11 seconds and 134.7 seconds, respectively, for the experimental group and 289.78 seconds and 93.0 seconds, respectively, for the control group. The observed difference was not significant, perhaps due to individual differences.

We were also interested in the difference between subjects' perception of reading time and their observed reading time. We calculated a residual time for each subject by subtracting their observed time from their own estimate. Of course, subjects' reading times varied, so we believe it is more meaningful to express their residuals as a proportion of their observed times. The Mann-Whitney U test⁴ showed that there was no significant difference between the residuals of the two groups ($U = 27, p > 0.05$), with the sum of the ranks equal to 99 for the experimental group and 72 for the control group.

With regards to the experiential statements, subjects in the experimental condition reported significantly less fatigue ($F[1, 6] = 10.000, p = .006$) and significantly greater enjoyment of the system ($F[1, 6] = 10.000, p = .006$) as compared to the control group. We followed up with median tests, which showed that for both the fatigue question ($p = .024$) and the system-enjoyment question ($p = .041$) there was a significant difference in the distributions of the two conditions. There was not a significant difference between the two groups with respect to their enjoyment of the story, the helpfulness of the projector, and how distracted they felt by the projector. Responses to the experiential statements are summarized in Table 4.

Eye Tracker Results

Subjects in the control group spent a greater percentage of their time looking at the LCD monitor ($M = 99.2\%$) than did the experimental group ($M = 95.7\%$). The Mann-Whitney U test⁵ showed a significant difference between the two groups' percentage of time looking at the monitor ($U = 6, p < .05$), with the sum of the ranks equal to 34 for the experimental group and 71 for the control group.

⁴ An ANOVA was not performed because the Hartley F-max test suggests that the assumption of homogeneity of variance may not hold for the reading time residuals. The critical F-max for $\alpha = .05, k = 2$, and $df = 8$ is 4.43 and our observed F-max was 6.78.

⁵ An ANOVA was not performed because the Hartley F-max test suggests that the assumption of homogeneity of variance may not hold for the percentage of time spent looking at the monitor. Our observed F-max of 8.30 exceeded the critical F-max of 5.82 ($\alpha = .05, k = 2, df = 6$).

Line of Sight	Condition		p-value
	Experimental	Control	
LCD Monitor	95.7%	99.2%	< .05*
Projection	4.1%	0.6%	< .05*
Other	0.2%	0.3%	.899

Table 2: Mean percentages of time spent looking at the LCD monitor, the projection, and other (content-free) areas

A corollary of this result was that the experimental group spent a greater percentage of their time looking at the projection on the wall ($M = 4.1\%$) than did the control group ($M = 0.6\%$). The Mann-Whitney U test⁶ showed a significant difference between the two groups' percentage of time looking at the monitor ($U = 4, p < .05$), with the sum of the ranks equal to 73 for the experimental group and 32 for the control group.

There was no significant difference in the percentage of time that the two groups spent looking at a content-free area of the room (anywhere other than the monitor or the projection): $F(1, 12) = 0.017, p = .899$. Table 2 summarizes the eye tracker results.

For our eye tracker, we experienced a sampling rate that ranged from 20 to 60Hz. In the analysis of the eye data, we ignored those samples that were recorded while a subject was blinking since obviously he or she was not taking in any visual information at that time. We also ignored the two samples immediately before and after a blink, because the eye tracker reported inaccurate data in those instances. For the average subject, these samples of blinks and "near blinks" amounted to less than 5%. In other words, for the average subject, over 95% of the recorded eye data were clean enough for analysis.

⁶ An ANOVA was not performed because the Hartley F-max test suggests that the assumption of homogeneity of variance may not hold for the percentage of time spent looking at the projection. Our observed F-max of 10.43 exceeded the critical F-max of 5.82.

Questions	Experimental		Control		ANOVA		Median Test
	M	SD	M	SD	F(1,16)	p-value	p-value
Picture	6.67	2.35	3.78	1.92	8.17	0.011*	0.024*
Reinforced	6.22	1.30	6.22	1.92	0.00	1.000	N/A
Text	12.11	4.04	11.33	3.57	0.19	0.671	N/A

Table 3: Number of Correctly Answered Memory Test Questions

Statement	Experimental		Control		ANOVA		Median Test
	M	SD	M	SD	F(1,16)	p-value	p-value
You enjoyed reading using this system.	5.33	0.71	3.67	1.41	10.00	0.006*	0.041*
You found the projector to be helpful.	4.00	2.06	3.44	1.01	00.53	0.479	0.103
You found the projector to be distracting.	3.44	1.42	3.67	1.66	00.09	0.764	0.143
You felt fatigued after using this system.	2.67	1.66	5.00	1.50	09.80	0.006*	0.024*
You enjoyed the story itself.	5.22	1.48	3.89	1.90	02.76	0.116	0.069

Table 4: Responses to Experiential Statements on a 7-point Likert Scale (1: strongly disagree ... 7: strongly agree)

DISCUSSION

The results show that the peripheral images did indeed convey additional information “for free” – that is, without adversely affecting comprehension of the text or reading time.

Subjects’ performance on the picture questions indicates that the images do communicate additional information. The absolute difference between the means of the two groups is 2.89 questions, which is an improvement of over 22%. Subjects in the control group performed better than we had expected on the picture questions, probably due to successful guessing. Despite the control group’s relatively impressive performance, both the ANOVA and the median test were significant.

The median test provides additional evidence that the peripheral border display system conveys more information than just the text alone. Non-parametric statistics require fewer assumptions than ANOVAs, but this flexibility comes at the price of committing more Type II errors (retaining the null hypothesis in the presence of a true effect). Although the median test might be employed in situations where data appear to violate the assumption of homogeneity of variance, this was not the case with our picture-question data ($F_{\max}[2,8] = 1.489, p < .05$). Here

we performed the median test only in order to reaffirm the result of the ANOVA.

From just this first result, it is still possible that the experimental subjects attended to the projected images to the detriment of their reading the text itself. However, since there was not a significant difference in performance on the text and reinforced questions, we conclude that peripheral images, in fact, did not adversely affect readers’ comprehension of the text.

These first two results suggest *that* the experimental group received more information with no text-comprehension penalty, but it does not indicate *how* they received it. From these two results alone, we could not rule out the possibility that experimental subjects studied each projected image with their foveal vision, incurring a cost of additional reading time. We would expect that if experimental subjects received the extra information through explicit examination of the images, then their reading times would be different from those of the control group. However, there was not a significant difference in reading time between the groups, so the extra information was received and retained without increasing task time.

Eye Tracker Discussion

The difference in time spent looking at the projection was very slight (only 3.5%), though statistically significant. We believe that a statistically genuine difference of such small magnitude is a low cost relative to the gain of additional information (22% improvement). The minimal effect of the eye data difference is reinforced by the fact that the reading times were not statistically different.

An examination of two experimental subjects suggests that the additional information found in the pictures can be obtained almost exclusively through peripheral vision. The two subjects who spent the smallest percentage of their time looking at the projection (0.72% and 0.99%) actually scored extremely well on the picture questions (correctly answering 10 and 8 questions, respectively). In fact, only one of the other 16 subjects scored better than these two on the picture questions, with a score of 10. The performance of these subjects suggests that additional information can indeed be obtained via peripheral vision.

RELATED WORK

Peripheral vision has been leveraged by other researchers to provide more immersive virtual-reality (VR) worlds [9] and to support multitasking [5]. Our use of peripheral vision is different from the former in that we use a desktop monitor and projector rather than a head-mounted display. The latter uses similar hardware to our system, but it is designed to facilitate multitasking, while our goal is to enhance a single reading task.

Baudisch et al. have combined a projector and LCD monitor to simulate a single, large, seamless display [1]. Our system uses a similar combination of hardware, but is different in that it uses the projector as a distinct display rather than merging it with the LCD monitor. Also, our technique employs a single, fixed focus area (the text on the LCD monitor), rather than a dynamic one. We intended that users never need to switch their focus away from the LCD monitor.

Tan et al. have used projected images to facilitate encoding and retrieval of primary content (lists of paired words) [11]. Although our use of projected images may incidentally reinforce information in the primary content, our main goal is to communicate entirely new information.

CONCLUSIONS AND FUTURE WORK

Our peripheral border display interaction technique showed how peripheral vision can be used as a basis to communicate additional information “for free” – without additional user effort and with no adverse effect on reading time or comprehension. We used an eye tracker to examine in greater depth how our interaction technique conveyed the additional information. Our lightweight, or non-command, interaction used projected images to convey environmental information in a situation where reading the text ought to remain the primary focus of the user’s attention.

There are several extensions of this system that might provide even greater benefit to users. The experience can be made richer by making the peripheral images more spatially immersive. For example, we could use a multi-projector system like CAVE™ [2] or use a single projector with a curved screen and reverse-distortion techniques similar to the Everywhere Display [8]. In addition to exploiting a visual channel, our system could be extended to convey even more information through an aural channel by playing background sounds that change as the user reads through the text. Just as the large peripheral images were blurred to avoid distraction, the sound would be greatly toned down and would change smoothly from one locale to another.

We could also incorporate eye tracking into the system to determine with greater precision where the user is reading [4, 12] in order to predict when to display each image. This technique would be especially helpful for text with more frequent picture changes than once per screen. More frequent imagery changes would grant the user even more “free” information.

Finally, our system could be used as an experimental tool to investigate whether users pay more attention to information in the text or in the periphery. The information that we presented in the images was congruent with the text, but our system could easily be modified to show conflicting information, and thereby determine which information source is dominant.

ACKNOWLEDGMENTS

We would like to thank our colleagues in the HCI Research Group at Tufts, including Jeffrey D’Amelia, Salem Berhanu, Julene Mazpule, and Jessica Webster. We also appreciate the assistance of Gregory Crane and the members of the Perseus Project in the Classics Department, and Richard Chechile of the Psychology Department. Furthermore, we would like to recognize Rikki Razdan and Alan Kielar from ISCAN for their timely technical support of the eye tracker. Finally, we are immensely grateful to our subjects who so generously volunteered for our study.

This research was principally supported by the National Science Foundation and National Endowment for the Humanities under the Digital Libraries Initiative Phase 2 Program. It was also supported in part by the NSF Interactive Systems Program, Office of Naval Research, Berger Family Fund, and the Tufts Selective Excellence Fund.

REFERENCES

1. Baudisch, P., Good, N., and Stewart, P. Focus Plus Context Screens: Combining Display Technology with Visualization Techniques. *Proc. UIST 2001*, ACM Press (2001), 31-40.
2. Cruz-Neira, C., Sandin, D., and DeFanti, T. Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. *Proc. SIGGRAPH 1993*, 135-142.

3. Jacob, R.J.K. Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces, in H. R. Hartson & D. Hix (Eds.), *Advances in Human-Computer Interaction, Vol. 4*, H.R. Hartson and D. Hix (Eds.), Ablex Publishing, Norwood, NJ, USA, 1993, 151-190.
4. Jacob, R.J.K. The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get, *ACM Transactions on Information Systems 10*, 3 (1991), 152-169.
5. MacIntyre, B., Mynatt, E.D., Volda, S., Hansen, K.M., Tullio J., and Corso, G.M. Support For Multitasking and Background Awareness Using Interactive Peripheral Displays. *Proc. UIST 2001*, ACM Press (2001), 41-50.
6. Murdocca, S. *Christmas Bear*. Simon and Schuster, New York, New York, USA, 1987
7. Nielsen, J. Noncommand User Interfaces. *Communications of the ACM 36*, 4 (1993), 83-99.
8. Pinhanez, C., Kjeldsen, R., Tang, L., Levas, A., Podlaseck, M., Sukaviriya, N., and Pingali, G. Creating Touch-Screens Anywhere with Interactive Projected Displays. *Proc. ACM Multimedia 2003*, ACM Press (2003), 460-461.
9. Robertson, G., Czerwinski, M., and van Dantzich, M. Immersion in Desktop Virtual Reality. *Proc. UIST 1997*, ACM Press (1997), 11-20.
10. Sibert, L. and Jacob, R.J.K. Evaluation of Eye Gaze Interaction. *Proc. CHI 2000, Conference on Human Factors in Computing Systems*, ACM Press (2000), 281-288.
11. Tan, D., Stefanucci, J., Proffitt, D., and Pausch, R. The Infocockpit: Providing Location and Place to Aid Human Memory. *Proc. Perceptive User Interfaces 2001*, ACM Press (2001).
12. Tanriverdi, V. and Jacob, R.J.K., Interacting with Eye Movements in Virtual Environments. *Proc. CHI 2000*, ACM Press (2000), 265-272.
13. New Human-computer Interaction Techniques for the Digital Library.
<http://www.cs.tufts.edu/~jacob/dl>.
14. Wickens, C. D., and Hollands, J. G. *Engineering Psychology and Human Performance*. 3rd ed. Prentice-Hall, Upper Saddle River, New Jersey, USA,, 2000