An Intrusive Evaluation of Peripheral Display

Xiaobin Shen*
School of Information Technology
University of Sydney

Andrew Vande Moere†
Key Centre of Design Computing and Cognitive
University of Sydney

Peter Eades‡
University of Sydney
National ICT Australia

Abstract

We present an intrusive evaluation of three different peripheral display systems on the same financial data set. Our results indicate there is a significant difference on a primary task performance and a peripheral comprehension task between large and small displays. Furthermore, we have found that distraction may be composed by *display-distraction* and *self-interruption*, and that animation may only influence the display-distraction. In addition, this paper proposes a measurement of *efficiency* derived from cognitive science. Finally, we propose three guidelines for peripheral display design.

CR Categories: B.4.2 [Input/Output and Data Communications]: Input/Output Devices—Image display; H.5.1 [Information Interfaces and Presentations]: Evaluation/Methodology—Screen design

Keywords: peripheral display, intrusive evaluation, large/small display

1 Introduction

The ambitious goal of peripheral display is to present information without distracting or overwhelming users. Obviously, it is hard to achieve this goal because there are two main limitations which block the development of peripheral display: limitations of human understanding and limitations of current technologies.

Our proposed research plans to conduct a systematic investigation of evaluation methods for peripheral display, including two parts:

Intrusive Evaluation, where the user is consciously aware of the evaluation experiment. This kind of evaluation often focuses on usability tests in a laboratory environment for a short period of time. Most often, such experiments are conducted using mainstream evaluation methods in information visualization (i.e. questionnaires, interviews and think aloud).

Non-Intrusive Evaluation, where the user is not consciously aware of the evaluation experiment. This often focuses on actual use in a general environment (*in situ*) for a long period of time, rather than a short-term laboratory situation. Currently, there are few successful methods that can be applied in this manner.

From subjects points of view, the biggest difference between intrusive and non-intrusive evaluation is mainly whether they know they are part of experiment or not. In intrusive evaluation, subjects actually realize that they are part of experiment but in a non-intrusive evaluation, they are not. Obviously, it is difficult to conduct an non-intrusive experiment because privacy issues must be considered in the experiment.

Here, we present the first part, which uses a laboratory controlled environment to evaluate three peripheral display systems in both large and small displays. Compared to non-intrusive evaluation, it has three advantages: it is easily conducted; it reveals results quickly and cheaply; it is easy to adopt existing evaluation methods and criteria to peripheral displays. Meanwhile, it has one big disadvantage in that results may lack ecological validity.

The general hypothesis for this experiment is that large displays are more effective in comprehension and efficiency with higher distraction than small ones. Our aim is twofold: to adapt a measurement of efficiency theory from cognitive science to measure the efficiency level in peripheral display; to discover the difference between large and small displays in the context of *distraction*, *comprehension* and *efficiency* .

2 Related Work

In this section, we will briefly review related work in the design and evaluation of peripheral displays.

2.1 Design

A number of projects have been designed and implemented in the field of peripheral displays. Most current peripheral display approaches use vision, audio and tactile senses. For example, InfoCanvas is visual example, which uses a beach scene to depict multiple pieces of real-time information [Plaue et al. 2004]. Non-Speech Audio Glance represents important properties of a message into a concise sound [Hudson and Smith 1996]. VisPad is a new haptic design for visualizing data, constructed from commodity massage chair pods with custom controllers and interfaces to a computer [Weissgerber et al. 2004].

It is difficult to have a good design and implementation in the field of peripheral displays because effective sound differentiation, motion perception and meaningful aesthetic consideration are all still open problems. Our experiment only uses vision as the information-carrying medium to implement peripheral display.

2.2 Evaluation

Although great progress has been achieved in the design of peripheral displays, few concentrate on the evaluation. However, there has been some work focusing on the evaluation of peripheral displays, which includes: Heuristic Evaluation, which adapted Nielsen's heuristic to the domain of peripheral displays [Mankoff et al. 2003]. Wizard of OZ, which simulates a wide range of plausible sensors to build multiple models [Hudson et al. 2003].

*e-mail: xshen@cs.usyd.edu.au
†e-mail:andrew@arch.usyd.edu.au
‡e-mail: peter.eades@nicta.com.au

Generally, current evaluation methods have limitations on suitability for the whole peripheral displays and in many cases, possibly ignore the real value of peripheral display. Here, our intrusive evaluation tries to capture the real value of peripheral display in the context of *distraction*, *comprehension* and *efficiency* instead of evaluation in general.

3 Peripheral Display Systems

Three systems are used in this experiment: Digital Data, MoneyTree and Stock Chart.

Digital data is a normal numerical visualization of financial data. It uses textual display to represent trade price and volume (see Figure 1(a)). It neither keeps previous trading history nor uses animation during transitions.

MoneyTree [Shen and Eades 2004] is a graphical display to represent financial data. It uses three different types of trees to represent three individual stock quotes and also uses tree length and leaves to represent the trade volume and price respectively (see Figure 1(b)) . No visualization history is depicted but image morphing is used to achieve smooth graphical transition from one stage to another.

Stock Chart is a traditional financial visualization, which uses three different colours to represent three stock quotes; a linear curve represents trade price and a bar chart represents trade volume (see Figure 1(c)). This shows trading history but there is no animation used in transition from one stage to another.

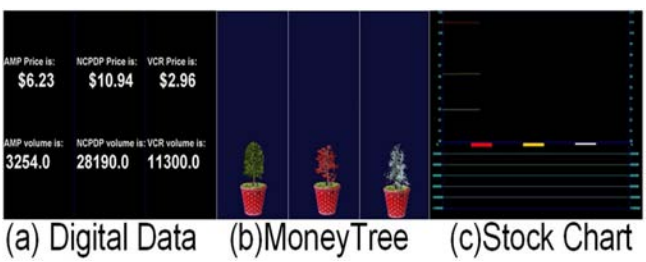

Figure 1: Three Financial Systems.

4 Experiment

The aim of this experiment was to discover the difference in the context of *distraction*, *comprehension* and *efficiency* between large and small displays.

Sixteen (8 female) subjects participated in our study, all having normal eyesight. Subjects ranged from 22 to 40 (the average is 28.6). 5 were master students, 9 were PhD students and 2 were postdoctoral researchers. Subjects were paid $20 each for their time.

Two standard 19 inch desk monitors were used in small displays and two large rear-projection screens used in large displays. Each of the large screens was 2m (length)*1.5m (height). We set the same resolution of 1024*768 in both experiments. We also adjusted the screen brightness and contrast to be identical. Figure 2 gives an illustration of the experimental setup.

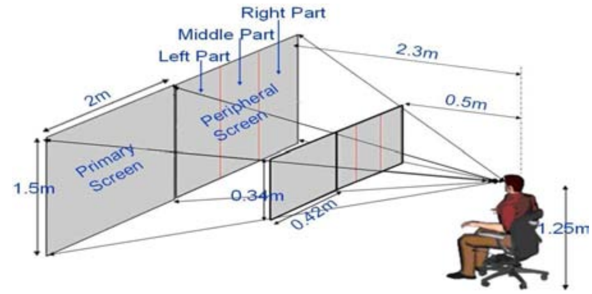

Figure 2: Experimental Setup.

One Dell Precision Workstation 360 and one LogiTech QuickCam Pro 4000 web camera were used in this experiment. The Dell Workstation was responsible for controlling the information displayed on both large and small displays. The web camera was used to count how many times subjects shift their focus to the peripheral screen by using OpenCV Face Detection [Intel 2005]. We also synchronize with the peripheral screen to save a snap shot of subjects during the content of peripheral display changing. By manually analysis of the snap shot image, we can calculate how many times subjects shift their focus to the peripheral screen because of peripheral display distraction. Then the rest of focus shifts subjects made is because of their own intention.

Desk, chair, keyboard and mouse were also involved in this experiment. The desk, chair and keyboard were fixed and the mouse was removed from the desk. A reference sheet for the three visualization metaphors was sitting in front of the keyboard in case subjects forgot the visual mappings.

Two hours trading from Australian financial stock information are used as data for all three systems. We also arbitrarily chose three companies stocks as our detailed display data source: AMP, NCPDP and VCR. The general trend for these three stocks was smoothly increasing for AMP and smoothly decreasing for NCPDP and VCR.

In each peripheral system, two tasks were involved: primary task and secondary task. The primary task was to allow subjects to type words on the primary screen for two minutes while the peripheral system was displayed on the peripheral screen. The typed words were from Shakespeare's Hamlet [Shakespeare 2005] and were printed on a one side A4 size sheet paper with a 16 point font. The secondary task was to occasionally shift focus to the peripheral screen to get information about the three stock quotes.

Three detailed hypotheses for this experiment was proposed: H1: Large display performs better than small display in terms of comprehension; H2: Large display achieves higher efficiency than small display; H3: The more times subjects shift their focus to the peripheral screen, the more information they will get.

A within-subject experimental design was used with the non-fixed ordering of experimental tasks. Once subjects signed the consent from, two tasks were given before conducting the actual trials: an introduction task and a practice task.

During the introduction task, each subject was given orientation material which included a definition of peripheral display, a brief introduction to peripheral display, a general information about three financial visualizations and sample questions and answers.

The last step before the actual trial was the practice task which was conducted on both large and small displays. Subjects were exposed to what the actual trials would be like. Within each experiment, there were four sub-tests (primary task plus Digital Data, primary task plus MoneyTree, primary task plus Stock Char and primary task without peripheral display).

5 Results

A paired student's t-Tests was used to analyze the results and a Z Score was used to calculate the efficiency level [Siegel and Morgan 1996].

Three independent parameters were analyzed: Mean Comprehension Error Rate (MCER), Mean Word Typed (MWT) and Mean Focus Shift (MFS). We use MCER to measure the *comprehension* in each test; use of MFS to measure the *distraction* and use of a combination of MCER, MWT and MFS to measure the *efficiency*.

The computation of efficiency was based on Pass and Merriënboer's efficiency measurement theory [Pass and Merriënboer 1993] from cognitive science. This theory proposes that efficiency is the combination of the measurements of mental effort and performance. We take the rough and ready view that MWT measures performance, MCER and MFS together measure the mental effort. All parameters are calculated in Z Score value. This measure is relatively simplistic, but it is adequate for our purpose.

$$E = \frac{Z_{MWT} - (Z_{MCER} + Z_{MFS})}{\sqrt{2}} \quad (1)$$

Our results reveal that large displays have a higher MCER than small displays and this difference is significant (See details in Table 1, where significant difference is $p<0.05$).

MCER	Test1	Test2	Test3
Large Display	3.53	2.94	2.54
Small Display	2.67	1.99	1.84
p	0.050	0.012	0.045

Table 1: Paired student's t-Test on MCER

Finer analysis of the results of Table 1 shows an unexpected result that MCER increases from left to right part on both large and small displays (See Figure 5) in the peripheral screen, which was divided into 3 parts, as shown in Figure 2.

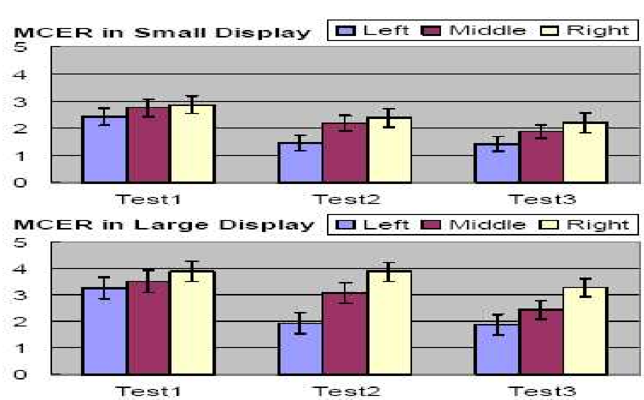

Figure 3: MCER in both small and large displays with 95% Confidence Interval

	Test1	Test2	Test3	Test4
t	-3.63	-2.18	-2.14	-2.08
p	0.003	0.048	0.049	0.998

Table 2: Paired student's t-Test on MWT

Table 2 reviews that large displays achieve lower MWT than small displays in all three tests. Also, this difference is significant($p<0.05$).

As we discussed in equation 1, efficiency is measured by combining consideration of performance and mental effort. Our results show that large displays have lower efficiency than small displays but the result is not significant; this needs further study.

6 Discussion

In this section, a brief discussion will be conducted in regarding to three hypotheses after we finished analyzing results.

6.1 Comprehension

As we mentioned in Results Section, comprehension in peripheral display is affected by relative position of the information in the peripheral screen. The farther the information is from the primary screen, the higher the comprehension error rate is (see Figure 2 for details).

In our experiment, we divide the peripheral screen which is next to the primary screen into three parts (See Figure 2). The viewing angle to the left part (around 35°) of peripheral screen is less than that to the middle (around 44°) and right part (around 53°). In the same way, the result of MCER to the left part is less than that to middle and right part in both large and small display experiments.

This result is in agreement with the theory of human vision. The biggest characteristic of human vision is the acuity outside of the fovea vision drops rapidly [Peripheral Version].

Conclusion 1: The most important information in the peripheral display has to be in the close proximity to the primary screen. This finding can be applied in many tasks (i.e. high-demand task, time critical tasks).

6.2 Distraction

Results from Section 5 also shows that animation leads to a significant difference in distraction between large and small displays. Here, we want to further divide distraction into two parts: display distraction and self-interruption.

Display distraction is caused by display itself and it can be divided into levels based on different data and tasks. Self-interruption occurs when subjects decide to look at the peripheral display themselves. Self-interruption is mostly affected by the characteristics of subjects themselves (personality, curiosity). It is difficult to decrease the level of distraction by self-interruption because this is under the control of subjects. However, the results in Section 5 indicate that it is possible to decrease the display distraction by using slow or smooth animation. This result also agrees with Bartram's finding that small periodic motions are better than color or shape cues especially in the periphery [Bartram 2001].

Conclusion 2: Slow, smooth animation can help to reduce the level of display distraction but may not affect the issue of the user's self-interruption.

6.3 Efficiency

Efficiency is one of the key concepts in the evaluation of normal information visualizations, but it is seldom used in peripheral displays. Here, our main purpose is to measure the task efficiency in peripheral displays.

In our experiment, the efficiency result in Test 3 is higher than that in Test 2 in both large and small displays because Test 3 has a higher performance and lower mental effort than Test2. There is one possible explanation for the lower mental effort that is previously known visual language. Most subjects have knowledge of stock chart (Test3) but few have knowledge of MoneyTree (Test2). Similarly, there is one possible reason to explain the higher performance. Subjects can mainly concentrate on the primary task in stock chart (Test3) than in MoneyTree (Test2) because stock chart (Test3) keeps previous visualization history.

Conclusion 3: considering previously known visual language and keeps previous visualization history can improve the efficiency of peripheral display.

6.4 Dsiplay Size

As we mentioned in Section 5, there is a significant difference in primary typing performance task (MWT) and peripheral comprehension task (MCER) between large and small displays.

There are two possible reasons to explain the significant difference in primary typing performance task. Firstly, subjects may take longer time to refocus from the typing sheet to the large screen than that to the small one. Secondly, it may be easier to detect the change in large screen than small one for the same display.

However, the size of screen is not the reason for causing significant difference in primary typing performance. As we mentioned in Section 5, the result of Test 4 (primary task without peripheral display) shows no difference.

Thus, we may draw a rough conclusion that small display performs better than large one, if there is a peripheral screen; if not, there is no difference. This result seems to be at odds with the results of Tan [Tan et al. 2003], which he shows that large display achieves better primary task performance than small one. A reasonable explanation may because Tan uses different kinds of data and requires pure focus.

Our result in Section 5 also reveals that there is a significant difference in MCER between large and small displays. This result also extends Tan's result [Tan et al. 2003] which he find there is no significant difference between large and small displays.

7 Conclusion and future work

In this paper, we present an intrusive evaluation comparing large and small display experiments for three peripheral display systems. From these experiments, we have two overall results: firstly,we put forward a measurement of efficiency from cognitive science into peripheral display. Secondly, we proposed three guidelines for peripheral display design:

- **G1**: The most important information in the peripheral screen should be close proximity to the primary screen.
- **G2**: Slow or smooth animation needs to be considered in the design of lower display distraction but may not affect the issue of the user's self-interruption.
- **G3**: Previous known visual language and keeps previous visualization history can enhance efficiency of peripheral display.

Clearly, a number of potential directions for follow-on work exist, including: to conduct a non-intrusive evaluation; to further investigate in the non-intrusive evaluation.

References

BARTRAM, L. 2001. *Enhancing Information Visualization with Motion*. Phd thesis, School of Computing Science, Simon Fraser Unversity.

HUDSON, S., AND SMITH, I. 1996. Electronic mail previews using non-speech audio. In *Proceedings of the conference companion on human factors in computing systems: common ground*, ACM Press, 237–238.

HUDSON, S., FOGARTY, J., ATKESON, C., AVRAHAMI, D., FORLIZZI, J., KIESLER, S., LEE, J., AND YANG, J. 2003. Predicting human interruptibility with sensors: A wizard of oz feasibility study. In *Proceeding of the conference on Human Factors in Computing Systems*, ACM Press, 257–264.

INTEL, 2005. Face detection. http://www.intel.com/cd/software/products/asmona/eng/219767.htm.

MANKOFF, J., DEY, A., HSIEH, G., KIENTZ, J., LEDERER, S., AND AMES, M. 2003. Heuristic evaluation of ambient displys. In *Proceeding of the conference on Human factors in computing systems*, ACM Press, 169–176.

PASS, F., AND MERRIËNBOER, J. 1993. The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors 35*, 737–743.

Peripheral version. http://webexhibits.org/colorart/ag.html.

PLAUE, C., MILLER, T., AND STASKO, J. 2004. Is a picture worth a thousand words? an evaluation of information awareness displays. In *Proceedings of the Graphics Interface*, Human-Computer Communications Society, 117–126.

SHAKESPEARE, W., 2005. Hamlet. http://www.pathguy.com/hamlet.htm.

SHEN, X., AND EADES, P. 2004. Using moneytree to represent financial data. In *Proceeding of eighth International Conference on Information Visualization*, IEEE Computer Society, 285–289.

SIEGEL, A., AND MORGAN, C. 1996. *Statistics and data analysis: an introduction*. John Wiley& Sons.

TAN, D., GERGLE, D., SCUPELLI, P., AND PAUSCH, R. 2003. With similar visual angles, larger displays improved spatial performance. In *Proceeding of the SIGCHI conference on human factors in computing systems*, ACM Press, 217–224.

WEISSGERBER, D., BRIDGEMAN, B., AND PANG, A. 2004. Feel the information with vispad: a large area vibrotactile device. *Information Visualization 3*, 1, 36–48.