Colour and Transparency on the Multi-Layer Display (MLD^{TM})

Anna Nees[1], Rochelle Villanueva[1] and Professor B.L. William Wong[2]

[1] University of Otago, Department of Information Science, Dunedin, New Zealand
neean240@student.otago.ac.nz,
RVillanueva@infoscience.otago.ac.nz
[2] Middlesex University, Interaction Design Center, School of Computing Science, London
w.wong@mdx.ac.uk

Abstract. It is a standard aim to complete tasks efficiently and effectively. When technology is involved, the tools must be designed to facilitate optimal performance. The ActualDepth^{TM} Multi-Layer Display (MLD^{TM}) is a 'new generation' display, consisting of two layered Liquid Crystal Displays (LCDs), with a gap between them. The top LCD displays transparently, allowing both layers to be viewed simultaneously. This paper describes an experiment that investigated relative reading speeds, error detection, comprehension speeds and comprehension accuracy on the MLD^{TM}, including a comparison with standard single layered displays. A framework pertaining to colour and transparency usage on the MLD^{TM} was then developed, which is intended to enhance the usability and effectiveness of the display. In general, it was found that overall readability was improved on the MLD^{TM}, compared to a standard display, and different transparency levels and colours should be employed, depending on the purpose of reading the text.

1 Introduction

This paper reports on an experiment conducted to determine which colour combinations and transparency levels support optimal reading performance on the MLD^{TM}, in its current state. The study also includes a comparison between reading performance on the MLD^{TM}, and on a standard single layered display.

The MLD^{TM} has the potential to increase information uptake, by taking advantage of human parallel processing. It is claimed that the MLD^{TM} will increase productivity by up to forty percent [10]. This is particularly beneficial in situations where the user is presented with information in an intense fashion, such as command and control systems [2]. Consequences of errors within system usage vary in gravity, from poor user performance, to possible loss of life. It is therefore important to research the MLD^{TM} in its current state, and propose a framework to enhance usability and interaction. The elements of colour and transparency are an integral part of the MLD^{TM}, and it is therefore appropriate that these issues are investigated, in order to support optimal performance.

To date, three basic guidelines for colour and transparency usage within the MLD™ have been devised [3]:

- The two screens of the MLD™ are lit from the back, so dark content placed on the rear screen will obscure the content on the front screen.
- With respect to transparency, noise and texture distracts the user's eyes from a layer, while drawing attention to the noise.
- Combining colours within the MLD™ is similar to overlaying two sheets of cellophane, with a light placed behind. It has been identified that by placing a darker colour over a lighter colour the resulting colour will not be as transparent as it would be with the lighter colour placed on top.

Little research has been done with regard to the usability the MLD™. However research regarding single layered displays has been performed, such as readability of text using different colour combinations. Previous research has shown that complementary colours cause undesirable flickering, resulting in eyestrain [5]; contrasting colours should be used for foreground/background combinations [6], [7], [8], [12], [14]; and negative polarity (dark text on light background) improves readability [11], [14], [15].

Previous research investigating transparent text [16] concluded that employing a textured background slowed reading times, and the more transparent the text was, the slower the reading times were.

In this experiment, it was hypothesized that previous findings mentioned above (pertaining to the standard single layered display), would also pertain to the MLD™.

2 Method

As the objective of the research was focused towards investigating colour combinations and transparency levels on the MLD™ in its current state, and the scope was narrowed to examine the task of reading text, the following factors of readability were tested:

i. Reading speed: how quickly a textual passage could be read.
ii. Error detection: how efficiently simple spelling errors could be recognized.
iii. Comprehension speed: how quickly text could be comprehended and recalled.
iv. Comprehension accuracy: how accurately text would be comprehended and recalled.

2.1 Experimental Design

The study used a between and within subjects, 4 x 12 repeated measures experimental design, that tested aspects of reading performance against display type (transparency level) and colour combinations (text colour/background colour).

For the purpose of data collection and analysis, the type of display (normal single layered display and MLD™) and their transparency level were combined and conceptually treated as the "display"; resulting in four types of 'displays' being tested:

- *SLD:* Standard single layered display
- *MLD:* MLD™
- *MLD+:* MLD™ with 30% more transparency added
- *MLD++:* MLD™ with 70% more transparency added

The twelve colour combinations were selected from the results obtained from previous research [1], [9]. It was decided to test twelve combinations in order to keep the experimental design within the scope of the study. Four colour combinations found to be good were selected; four found to be bad, and four where research was inconclusive as to their effectiveness. These colours were selected to represent those commonly used, and to ensure that a variety of colours were tested.

The twelve colour combinations were divided into two groups (A and B), so each participant was only exposed to six of the twelve combinations. However within a single experiment, each combination was tested twice (to ensure consistency), and hence each participant was presented with twelve experimental trials. The assignment of the colour combinations was performed randomly, to eliminate opportunity for participant prediction. The twelve selected colour combinations, together with their group and trial allocations are shown in the table below:

Table 1. Colour combinations selected from previous research findings [1], [9]

	Assigned Group	Trials	Text Colour	Background Colour
Good Combinations	A	T5, T12	Black	White
		T1, T10	Magenta	Blue
	B	T5, T12	Green	Black
		T1, T10	Yellow	Blue
Poor Combinations	A	T3, T8	Green	White
		T6, T11	Blue	Red
	B	T3, T8	Blue	Black
		T6, T11	Red	Green
Effectiveness Unknown	A	T2, T9	Light Magenta	Green
		T4, T7	Yellow	Magenta
	B	T2, T9	Green	Red
		T4, T7	Yellow	Cyan

Table 2 summarizes the transparency levels and display types employed across the two colour groups.

Table 2. Transparency levels and displays tested within the two colour groups

	Colour Group	Applied Transparency Level (%)
MLD™	A	0* (MLD)
		30 (MLD+)
		70 (MLD++)
	B	0* (MLD)
		30 (MLD+)
		70 (MLD++)
SLD	A	0 (SLD)
	B	0 (SLD)

** "0%" transparency on the MLD™ equates to the standard transparency on the front screen, with no further transparency being applied*

A practice trial was also undertaken, to familiarize participants with the experiment. This practice used black text on a yellow background; and was chosen as to not 'skew' the results, by giving a preview of any colour combinations in the experiment. The practice could be repeated as many times as desired, until the participant felt confident to continue with the actual experiment.

2.2 Equipment

Four experimental systems were developed, using VisualBasic 6.0 (one for each colour group for each of the two displays). Systems designed for the single layered display were implemented with the text overlaid on the background, and for the screen dimensions of 1024 x 768 pixels. Systems designed for the MLD™ were twice the width (2048 x 768 pixels), to reflect the way the MLD™ treats the display as a 'double-width' standard display, that 'wraps' the content around. Therefore text and buttons were put on the left hand side, and the background on the right hand side. Participants navigated through the system by clicking on buttons, which also acted as a trigger to start and stop the timers (for each reading and set of comprehension questions).

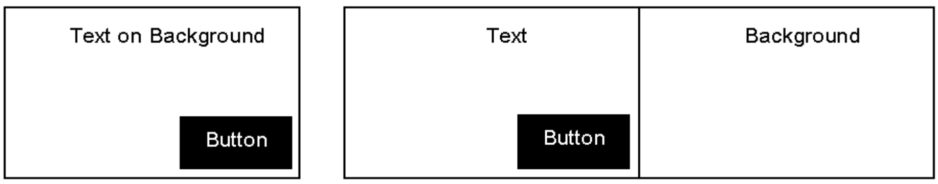


Fig. 1. System design for the standard single layered display (left) and the MLD™ (right)

All texts were left justified, with single line spacing, and one line spacing between paragraphs. The font was Tahoma, size 10 point, as this font was found to support optimal reading performance on a screen in previous research [13]. All buttons were in the bottom right hand corner, and all screens with readings included a consistent 'textured' background, to reflect the typical use of the MLD™.

Two identical computers were used; one for each of the display types. Both displays were 15 inches, and accompanied with a keyboard and mouse.

2.3 Participants

Forty participants took part in the experiments. All participants were eighteen years or over; spoke English as a first language; had good or corrected eyesight and had common trichromat vision (normal colour vision). Participants were randomly divided into the groups, with five participants being tested for a specific colour set and display (and transparency) type.

2.4 Readings and Tasks

Readings were evaluated using the Flesch reading difficulty index [4], to ensure consistency. The level of difficulty of all readings used had a readability index of between sixty and sixty-five (classified as "plain English", where zero is "very difficult" and one hundred is "very easy"). Each reading was divided into two parts of approximately 150 words, with a line length of approximately 70 characters. This was to avoid the need for scrolling or paging, which would have introduced new variables.

Each reading contained either six or eight spelling errors. As participants were exposed to each colour combination twice, one instance of each combination contained six errors, and the other contained eight. All spelling errors were simple and common words, to ensure the ability to *detect* errors was being tested (not the knowledge that a word was spelt incorrectly). An implemented counter recorded the number of errors detected for each reading, which was triggered by the participant pressing the 'Enter' key on the keyboard.

At the conclusion of each reading, participants were presented with a set of five comprehension multiple choice questions about the reading.

In order to apply transparency to the MLD™, a tool called 'Vitrate' was employed. This is a utility for Windows 2000 (or newer) that allows the transparency of windows to be adjusted, between zero and ten.

3 Results and Discussion

The system used in the experiments automatically recorded the times taken to complete each reading, and each set of comprehension questions; the number of errors detected in each reading; and the answers given for comprehension questions.

After removing any values greater than two standard deviations from the mean (to prevent the results being skewed), a repeated measures ANOVA was used, to determine whether or not differences in the results were significant.

3.1 Transparency and Display Type

The following figures illustrate the mean results for each of the types of readability tested, between the groups, showing the variances across the different display types and transparency levels.

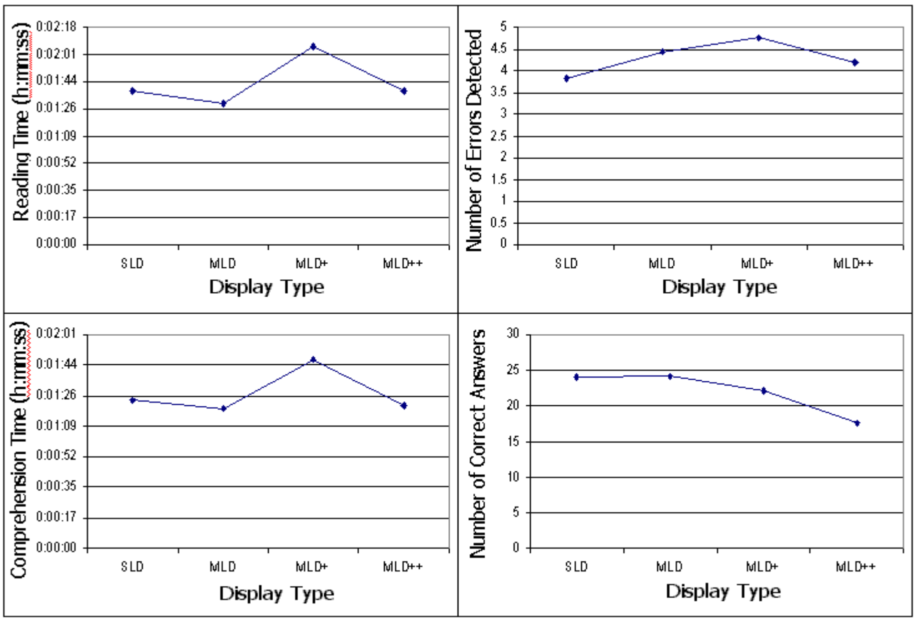


Fig. 2 Mean results for reading times (top left), number of errors detected (top right), comprehension times (bottom left) and number of correct comprehensions (bottom right) across the display types and transparency levels

Key Findings Results from the repeated measures multi-factor ANOVA show that there was a significant display effect for the tested readability types (reading speed: ($F_{(3, 76)} = 7.380$, $p = 0.000$), error detection: ($F_{(3, 76)} = 2.060$, $p = 0.001$), comprehension speed: ($F_{(3, 76)} = 5.912$, $p = 0.001$) and comprehension accuracy: ($F_{(3, 76)} = 10.702$, $p = 0.000$)). Results regarding display type and transparency are summarized below:

i. The MLD™ in its current state supported the best-observed performance with respect to reading speed, and comprehension speed and accuracy.
ii. Improved error detection resulted from 30% transparency being added to the MLD™. As this display type did not perform well with respect to other aspects of readability, it can be assumed that this level of transparency slows the reading process down, making errors more apparent to the reader.
iii. With the exception of error detection, reading performance was improved on the MLD™ with 70% transparency, compared to when just 30% transparency was added, however the reasoning for this is not apparent.
iv. Reading speeds and comprehension speeds were increased, and comprehension accuracy was improved on the standard single layered display, compared

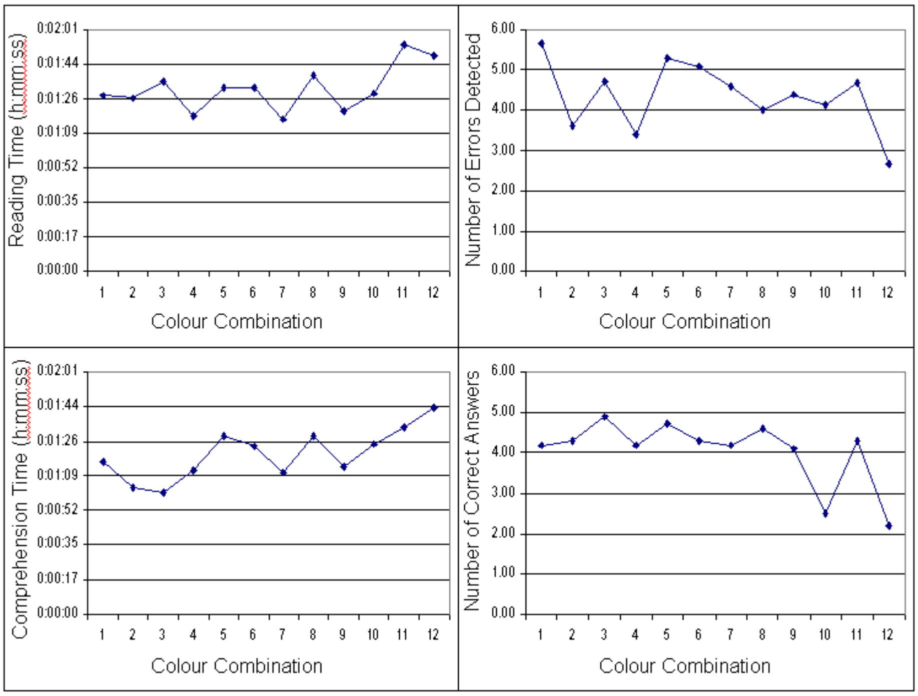
to the MLD™ with 30% and 70% transparency added; however demonstrated the poorest error detection performance; indicating that errors are more obvious when the text is projected out from the background.

v. Comprehension accuracy performance decreased in a near-linear fashion, as the transparency on the MLD™ increased, indicating that higher transparency levels result in text that is harder to comprehend, presumably due to concentration being devoted to discerning the actual words.

vi. Comparing the best and the poorest performance on the displays, there was approximately a 40% difference in reading speeds; 25% difference in error detection; 40% difference in comprehension speeds; and 40% difference in comprehension accuracy.

While it appears that, with the exception of error detection, readability is well supported by the MLD™ in its current state, it would be desirable to test further levels of transparency at smaller intervals, to confirm this finding. This would allow for specific usage recommendations to be made, as opposed to very generalized guidelines.

3.2 Colour Combinations



1. Black on White	2. Yellow on Blue	3. Magenta on Blue	4. Green on Black
5. Green on White	6. Blue on Red	7. Red on Green	8. Blue on Black
9. Green on Red	10. Yellow on Cyan	11. Light Magenta on Green	12. Yellow on Magenta

Fig. 2 Mean results for reading times (top left), number of errors detected (top right), comprehension times (bottom left) and number of correct comprehensions (bottom right) across the colour combinations

Figure 2 illustrates the mean results for colour combinations employed on the MLD™ in its current state. On each graph, the first four colour combinations along the x-axis are those previously found to be good; the middle four are those found to be poor, and the right four those where their effectiveness is unknown ([1], [9]).

Key Findings Results from the repeated measures multi-factor ANOVA show that there was a significant colour combination effect for the tested readability types (reading speed: ($F_{(5, 180)} = 4.975$, $p = 0.000$), error detection: ($F_{(5, 190)} = 3.006$, $p = 0.012$), comprehension speed: ($F_{(5, 190)} = 20.653$, $p = 0.000$), and comprehension accuracy: ($F_{(5, 360)} = 48.414$, $p = 0.000$)). Results regarding display type and transparency are summarized below:

i. For each aspect of readability tested, there is no apparent correlation between the determined effectiveness of various colour combinations in previous research ([1], [9]), and their associated performance on the MLD™. This demonstrates that there are usability differences between a standard display and the MLD™, and emphasizes the need to reassess colour combination recommendations for the MLD™, rather than applying the same rules as recommended for standard displays.

ii. Colour combinations of high saturation (such as magenta and blue) slow reading on the MLD™ down.

iii. Colours of low contrast (such as blue and black) result in poor reading performance, with all four types of readability tested. This correlates with previous findings and recommendations ([6], [7], [8], [12], [14]).

iv. Complementary colours (such as yellow and blue) are not necessarily poor with respect to reading speed, and comprehension speeds. This conflicts with previous findings [5], which suggests that complementary colours cause eye flicker. However complementary colours degraded performance with respect to error detection and comprehension accuracy. This indicates that the 'eye flicker' thought to result from such combinations may in fact speed processes such as reading and comprehension recall, however degrades tasks requiring attention to detail, such as error detection, and recalling the specifics of the content.

v. Negative (dark on light) polarity combinations (such as green on white and black on white) supported improved error detection, which agrees with previous findings that negative polarity improves readability ([11], [14], [15]). However combinations of positive polarity (light on dark) were seen to best facilitate faster comprehension speeds and comprehension accuracy. While a reason for this is not certain, it is possible that darker text allows the reader to focus on the actual characters of the text, facilitating error detection, while lighter text does not make individual characters stand out so much, allowing the general flow to text to be read faster.

vii. Comparing the best and the poorest performance associated with colour combinations, there was approximately a 50% difference in reading speeds; 110% difference in error detection; 70% difference in comprehension speeds; and 120% difference in comprehension accuracy.

While general recommendations with respect to colour combinations can be made from these findings, it would be desirable to extend this study by testing a greater range of combinations, as only testing twelve limits the degree to which specific recommendations can be made.

3.3 Transparency and Display Types and Colour Combinations

While there were significant transparency and display type effects, and significant colour combination effects, there ANOVA tests found there to be no significant interaction between these two, for any aspect of readability (reading speed: ($F_{(25,180)}$ = 1.232, p = 0.201), error detection: ($F_{(25,190)}$ = 1.065, p = 0.386), comprehension speed: ($F_{(25,190)}$ = 1.181, p = 0.262) and comprehension accuracy: ($F_{(25,190)}$ = 1.065, p = 0.386)).

4 Conclusions

In general, results from this experiment differ to those obtained from similar research with standard single layered displays, emphasizing the importance of reassessing design recommendations for the MLD™. Results from this study therefore do not support the hypothesis that design recommendations for standard displays will relate, and can also be applied to the MLD™.

Results have been obtained, from which recommendations can be made with regard to colour and transparency usage on the MLD™. No further transparency should be applied to the display, although it has been shown that further transparency can enhance tasks such as error detection. In general, colour combinations of high contrast should be employed, with low saturation. To ensure optimal readability, it is recommended that complementary colours be avoided (although these do not necessarily degrade performance, they also do not optimize it). For tasks requiring attention to be paid to detail, use dark on light polarity, however light on dark polarity should be used for tasks corresponding to the actual content of the passage (specifically comprehension tasks).

The results suggest that the colour combinations and transparency levels employed on the MLD™ can greatly affect reading performance. Readability can be affected by 25-40%, depending on the display type and transparency level used, and 50-120% depending on the colour combinations selected. However when making such design choices, it is important to keep in mind what the goal of the task is; fast reading; error detection; fast information recall; or accurate information recall, as not all these tasks are best supported by the same design specifics.

References

1. Bailey, R.W. and Bailey, K.N., "Protoscreens: Rapid Prototyping Software", Bailey & Bailey Software Corporation, Ogden, Utah, 1992
2. Bell, G., Engel, G., Evanicky, D., "Search for Change with Multi Layer Displays (MLD) – Applied Vision/Human Factors: Display Ergonomics and Standards", Deep Video Imaging Ltd, 2003
3. Deep Video Imaging, "Multi-Layer Display Technology Guide", Deep Video Imaging Ltd, 2003
4. Flesch, R., "How to Write Plain English" Chapter 2, Harper and Row, 1979
5. Fowler, S.L. and Stanwick, V.R. "The GUI Style Guide" Fast Consulting, Cambridge, MA: Academic Press, 1995
6. Galitz, W., "The Essential Guide to User Interface Design" – Chapter 12: "Choose the Proper Colors", 2nd Ed., John Wiley & Sons Inc., U.S.A., 2002
7. Hill, A. and Scharff, L. V. "Readability of screen displays with various foreground/background color combinations, font styles, and font types", Proceedings of the Eleventh National Conference on Undergraduate Research, 2,1997, pp. 742-746
8. Hill, A. L. and Scharff, L. V., "Readability of Computer Displays as a Function of Colour, Saturation, and Background Texture", D. Harris Ed., Engineering Psychology and Cognitive Ergonomics, Brookfield USA, Vol. 4, 1999, pp. 123 - 130
9. Lalomia, M., and Happ, A., "The Effective Use of Color for Text on the IBM 5153 Color Display", Proceedings of the Human Factors Society 31st Annual Meeting, 1987, pp. 1091 – 1095
10. Larson, B., "First Look: Actual Depth", Tech Live, accessed February 2002, http://www.techtv.com/products/hardware/story/0,23008,3374239,00.html
11. Lynch, P., Horton, S., "Web Style Guide: Basic Design Principles for Creating Web Sites", 2nd Ed., Yale University Print, 2002
12. Marcus, A., "Graphic Design for Electronic Documents and User Interfaces", ACM Press/Addison-Wesley Publishing Co., 1992
13. Milligan, P., Nees, A., Barltrop, K., Villanueva, R., Wong, W., "Arial, Georgia, Times or Verdana? Which is Better?", Unpublished Research Article, 2002
14. Parker, B., and Scharff, L. V., "Influences of Contrast Sensitivity on Text Readability in the Context of a GUI", http://hubel.sfasu.edu/research/agecontrast.html, accessed June 2003
15. Pastoor, S., "Legibility and Subjective Preference for Color Combinations in Text", Human Factors, 32(2), 1990, pp. 157 – 171
16. Scharff, L. F. V., and Ahumada, A. J., Jr., "Predicting the Readability of Transparent Text", Journal of Vision, Vol. 2, Iss., 9, 2002, pp. 653 – 666